

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 03:47:51 ; Search time 211 Seconds
(without alignments)
7816.905 Million cell updates/sec

Title: US-10-724-972A-2580

Perfect score: 1008

Sequence: 1 ggagtggaatcagtgagagg.....gtaaaaagataataagtaa 1008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	222.6	22.1	668	4 US-08-956-171E-355
2	222.6	22.1	668	4 US-08-781-986A-355
3	162.2	16.1	242	4 US-08-956-171E-2556
4	162.2	16.1	242	4 US-08-781-986A-2556
5	158.2	15.7	183	3 US-09-134-001C-2023
6	125.4	12.4	2115	4 US-08-956-171E-604
7	125.4	12.4	2115	4 US-08-781-986A-604
8	124	12.3	999	4 US-09-830-217-15
9	124	12.3	999	4 US-10-278-946-15
10	124	12.3	3775	4 US-08-956-171E-238
11	124	12.3	3775	4 US-08-781-986A-238
12	109.4	10.9	3046	4 US-09-710-279-3782
13	77.2	7.7	1056	4 US-09-543-681A-2690
14	73.8	7.3	421	4 US-08-956-171E-383
15	73.8	7.3	421	4 US-08-781-986A-383
16	72.2	7.2	612	4 US-09-902-540-1357
17	70.6	7.0	975	3 US-09-071-035-37
18	70.6	7.0	984	4 US-09-134-000C-937
19	69.8	6.9	709	4 US-09-902-540-1583
20	68.8	6.8	702	4 US-09-902-540-2208
21	68.8	6.8	7218	1 US-08-232-463-14
22	65.6	6.5	889	3 US-09-071-035-39
23	64.6	6.4	3153	4 US-09-710-279-3348
24	64.6	6.4	3267	4 US-09-710-279-4042
25	64.6	6.4	3618	4 US-09-710-279-3564
26	64.4	6.4	1044	4 US-09-710-279-461
27	64.4	6.4	1044	4 US-09-710-279-1267

Sequence 2655, Ap
Sequence 1280, Ap
Sequence 22, Appl
Sequence 1393, Ap
Sequence 37, Appl
Sequence 37, Appl
Sequence 75, Appl
Sequence 13, Appl
Sequence 1, Appl
Sequence 66, Appl
Sequence 14624, A
Sequence 1318, Ap
Sequence 1, Appl
Sequence 3335, Ap
Sequence 1, Appl
Sequence 1, Appl
Sequence 862, App

64.4 6.4 1056 3 US-09-134-001C-2655
63.6 6.3 1039 4 US-09-902-540-1280
62.2 6.2 1141 4 US-09-806-708B-22
62 6.2 978 4 US-09-543-681A-1393
61.8 6.1 9834 4 US-08-956-171E-37
61.8 6.1 9834 4 US-08-781-986A-37
59.8 5.9 5652 4 US-09-601-198-75
59 5.9 19124 2 US-08-487-826B-13
59 5.9 640681 4 US-09-790-988-1
58.8 5.8 861 4 US-09-949-016-14624
56.4 5.6 147382 4 US-09-902-540-1318
56.4 5.6 1664976 4 US-08-916-421B-1
56.4 5.6 1664976 4 US-09-692-570-1
55.2 5.5 4404 4 US-09-134-000C-3335
55.2 5.5 1956 3 US-08-559-896B-1
55 5.5 1956 4 US-09-351-794A-1
54.8 5.4 1059 4 US-09-107-532A-862

ALIGNMENTS

RESULT 1

US-08-956-171E-355
; Sequence 355, Application US/08956171E

; Patent No. 6593114

; GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 355:

SEQUENCE CHARACTERISTICS:

LENGTH: 668 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 355:

US-08-956-171E-355

Query Match 22.1%; Score 222.6; DB 4; Length 668;
Best Local Similarity 74.8%; Pred. No. 8.3e-39;
Matches 279; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 627 TTTCCTGCTCATCCAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAA 686
Db 1 TTTATTAGCACATCCAAACTATTTCATATGTTGGACAATTTTAAACGCACTAGGATTAA 60

QY 687 AGAAGCAATTAAGTGATGATGTTACTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTT 746
Db 61 AAATGCAATTAAGTGACGATGTAACAAAGGTTTAAAGTAAATATTTGAAAGGACCTTACTT 120

QY 747 ACAAATGAACACTGAACTTTATCTCAAGTGAATCCCTGAGGCTATGTTTCATAATGACAAA 806
Db 121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGGCTATGATCATTTATGACAGA 180

QY 807 CAAAGCAAGTTCTTAACGAACCTTCACTAAAGAACTAGAAAAGATCCGTATGGAAGAA 866
Db 181 TCATGCTAAAAGAAATTCCTGCTGAATTCAGAAAGTTACAGAGATGCAATGGAAGAA 240

QY 867 ATTAACGCTGTGAAAAATCAACGTTGTTGATATTTTAGACCGTGACTTTATGGGCAAGATC 926
Db 241 GTTGAATGCAGTTAAAATAATCGCTGATATTTGACCGTGATGTTTGGGCAAGATC 300

QY 927 ACGTGGTTAAATTTCTCAGAGAAATGCGCAAGAACTTGTGAAATTTCTAAGAAAGA 986
Db 301 TCGTGGCTTAATTTCTTCTGAAGAAATGCGCTAAGAACTTGTGAAATTTCAAAAAAGA 360

QY 987 TAGTAAAAAAGAT 999
Db 361 ACAAAAGTAAGGT 373

RESULT 2

US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

US-08-781-986A-355

Query Match 22.1%; Score 222.6; DB 4; Length 668;
Best Local Similarity 74.8%; Pred. No. 8.3e-39;
Matches 279; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 627 TTTCCTGCTCATCCAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAA 686
Db 1 TTTATTAGCACATCCAAACTATTTCATATGTTGGACAATTTTAAACGCACTAGGATTAA 60

QY 687 AGAAGCAATTAAGTGATGATGTTACTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTT 746
Db 61 AAATGCAATTAAGTGACGATGTAACAAAGGTTTAAAGTAAATATTTGAAAGGACCTTACTT 120

QY 747 ACAAATGAACACTGAACTTTATCTCAAGTGAATCCCTGAGGCTATGTTTCATAATGACAAA 806
Db 121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGGCTATGATCATTTATGACAGA 180

QY 807 CAAAGCAAGTTCTTAACGAACCTTCACTAAAGAACTAGAAAAGATCCGTATGGAAGAA 866
Db 181 TCATGCTAAAAGAAATTCCTGCTGAATTCAGAAAGTTACAGAGATGCAACATGGAAGAA 240

QY 867 ATTAACGCTGTGAAAAATCAACGTTGTTGATATTTTAGACCGTGACTTTATGGGCAAGATC 926
Db 241 GTTGAATGCAGTTAAAATAATCGCTGATATTTGACCGTGATGTTTGGGCAAGATC 300

QY 927 ACGTGGTTAAATTTCTCAGAGAAATGCGCAAGAACTTGTGAAATTTCTAAGAAAGA 986
Db 301 TCGTGGCTTAATTTCTTCTGAAGAAATGCGCTAAGAACTTGTGAAATTTCAAAAAAGA 360

QY 987 TAGTAAAAAAGAT 999
Db 361 ACAAAAGTAAGGT 373

RESULT 3

US-08-956-171B-2556
; Sequence 2556, Application US/08956171B
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2556:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2556:
US-08-956-171E-2556

Query Match 16.1%; Score 162.2; DB 4; Length 242;
Best Local Similarity 79.3%; Pred. No. 6.4e-26;
Matches 191; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 208 GATCGTTAGTTCCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATTAACAAAAAAT 267
DB 2 GATGCATTAGCAGCATTAGACGTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 61

QY 268 CGTATTATTAACCATTAAGATTAATAATTTGGAATACACTTCTGTAGGACACGTAAG 327
DB 62 CGTATCATTAAACCGATTAGACGTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 61

QY 328 CAACCTTAACCATTAAGATTAATAATTTGGAATACACTTCTGTAGGACACGTAAG 387
DB 122 CGTATCATTAAACCGATTAGACGTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 121

QY 388 AGACACAAAGGATTTATTAAGATTTAAATTAATAATTTGCTTACGATTTGAACTGAAAAAGT 447
DB 182 AGACATANAGGTATTATTAAGATTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 241

QY 448 T 448
DB 242 T 242

RESULT 4
US-08-781-986A-2556
Sequence 2556, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: FB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2556:
SEQUENCE CHARACTERISTICS:

LENGTH: 242 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-2556

Query Match 16.1%; Score 162.2; DB 4; Length 242;
Best Local Similarity 79.3%; Pred. No. 6.4e-26;
Matches 191; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 208 GATCGTTAGTTCCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATTAACAAAAAAT 267
DB 2 GATGCATTAGCAGCATTAGACGTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 61

QY 268 CGTATTATTAACCATTAAGATTAATAATTTGGAATACACTTCTGTAGGACACGTAAG 327
DB 62 CGTATCATTAAACCGATTAGACGTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 121

QY 328 CAACCTTAACCATTAAGATTAATAATTTGGAATACACTTCTGTAGGACACGTAAG 387
DB 122 CAGCCAAACTTTAGAAGAAATTTAGTAAATTTAAACCGGATTTAATTCGCTGATAGCAGT 181

QY 388 AGACACAAAGGATTTATTAAGATTTAAATTAATAATTTGCTTACGATTTGAACTGAAAAAGT 447
DB 182 AGACATANAGGTATTATTAAGATTTAAACCGATTGGTATTCGTGATGATGGTAAGAAAAA 241

QY 448 T 448
DB 242 T 242

RESULT 5
US-09-134-001C-2023
Sequence 2023, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2023
LENGTH: 183
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2023

Query Match 15.7%; Score 158.2; DB 3; Length 183;
Best Local Similarity 98.2%; Pred. No. 4.3e-25;
Matches 160; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 846 AAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTGA 905
DB 21 AAAAAGACCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTGA 80

QY 906 CCGTGACTTATGCGCAAGATCACGTGGTTTAAATTTCTCAGAAGAAATCGCAAAAGAACT 965
DB 81 CCGTGACTTATGCGCAAGATCACGTGGTTTAAATTTCTCAGAAGAAATCGCAAAAGAACT 140

QY 966 TGTGAAATTTCTAAGAAAGATAGTAAAGAAATTAAGTAA 1008
DB 141 TGTGAAATTTCTAAGAAAGATAGTAAAGAAATTAAGTAA 183

RESULT 6
US-08-956-171E-604
Sequence 604, Application US/08956171E

Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-08-956-171E-604
Query Match 12.4%; Score 125.4; DB 4; Length 2115;
Best Local Similarity 53.2%; Pred. No. 8.1e-18;
Matches 289; Conservative 1; Mismatches 247; Indels 6; Gaps 1;
QY 458 ATTATAATGAAATATTGATGCTTTTAAACAAATTTCAAAAGCTTTAGGTAAAGAAAG 517
DB 1 ATTATAATGCAATATTGAGCAATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAAG 60
QY 518 AAGTAAAGAAAGCTTTAGAGAACGATAGAAAATTTGAAGAAATATAAAAGAAATAA 577
DB 61 AAGCGAGAAGCGCTGGAAGAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAATTTG 120
QY 578 CTATGATGATAAAATCAAAAGGATTTCCTGCGAGTAGCTGCTAAATCAGGTTTGTCTC 637
DB 121 AACAGAGTACGTTAAATCTGCAATTTGCAATTCGCTATCTCAAGAGCAGGTATGTTTATA 180
QY 638 ATCCAAGCAACTTTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAAGCAATTA 697
DB 181 ATAATGAAGATACATTTATGGGACAAATCTTAAATTAATGGTATTCAACCTGAAGTCA 240
QY 698 GTGATGATGTTCTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTTACAATGAACA 757
DB 241 NAAAR 300
QY 758 CTGAACCTTATCTCAAGTGAATCTGCGGATGTTTCATATGATGATGATGATGATGATGAT 817

DB 301 ATGAAGAACTTGCCAATATCAATCAAAAGATTATGATTTTATGCCACTGACGGAAGAAACGG 360
QY 818 CTAACGAACTTCTACTAAAGAACTAGAAAAGATCCTGTATGGAAGAAATTAACGCTG 877
DB 361 ACAAAAA-----TAGAAGCAAAATTCATTGATCCTGCAAGTTTGGAAATCATTTAAAGCTG 414
QY 878 TGAATAATCAAGCTGTTGATATTTTAGACCTGACTTATGGGCAAGATCACTGTTGTTAA 937
DB 415 TGAAGATAACAAAGTTTATGACCTTGACCGAAATAGTGTGTAATCAAGGGGGATTA 474
QY 938 TTTCTTCAGAAAGAAATGCAAAAGAACTTGTGTAATTTATCTAAGAAAGATAGTAAAAAG 997
DB 475 TCGCAAGTGAAGTATGCGAGAAGATTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534
QY 998 ATA 1000
DB 535 ACA 537
RESULT 7
US-08-781-986A-604
Sequence 604, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-604

Query Match 12.4%; Score 125.4; DB 4; Length 2115;
Best Local Similarity 53.2%; Pred. No. 8.1e-18;
Matches 289; Conservative 1; Mismatches 247; Indels 6; Gaps 1;
QY 458 ATTATAATGAAATATTGATGCTTTTAAACAAATTTCAAAAGCTTTAGGTAAAGAAAG 517
DB 1 ATTATAATGCAATATTGAGCAATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAAG 60
QY 518 AAGTAAAGAAAGCTTTAGAGAACGATAGAAAATTTGAAGAAATATAAAAGAAATAA 577
DB 61 AAGCGAGAAGCGCTGGAAGAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAATTTG 120


```
Query Match      12.3%; Score 124; DB 4; Length 999;
Best Local Similarity 50.7%; Pred. No. 1.4e-17;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

Qy 27 AATTTTAAAGTGAATGGCTTATTGTTTAAATGCAACTGCAGCATGTGGAAATAA 86
Db 18 AATTTAAATGCTTGTGTAGCTTGTCTTCTACTTGTTTTAGCAGGATGTAGTGGAA 77

Qy 87 TAGTTCAAGTAACTCAAGTAAGAGTATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
Db 78 TTCAAATAAACAATCATCTGATAACAAGATAAGGAAACAACCTTCAATTAACAATGCAAT 137

Qy 147 AGGTACTACGAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGTAGTATTCAATTGT 206
Db 138 GGGTACAACTGAAATTAAGGGAACCAACGCGTGTGTTACGCTATATCAAGGTGCCAC 197

Qy 207 TGATGCGTTAGTGTCTTTAGATGTTTAAACCTGTTGGGATAGCGGATGATAACAATAA 266
Db 198 TGACGTGCTGTATCTTTAGTGTGTTAAACCTGTAGTGTCTGTAGNATCATGACACAAA 257

Qy 267 TCGTATTATTAACCATTAAGAGATAAAATTTGGAATAATACACTTCTGTAGGAACACGPA 326
Db 258 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATACTAAGATTGTAGGTCAAGAAC 317

Qy 327 GCAACCTAACTAGNAGGAATCAGTAATCTTAAACCCAGATTTAATTATGCTGATATAA 386
Db 318 TGCACCTAACTAGAGGAATCTCTAAATTAACCCGCGACTTAAATGTCGCGTCAAAAGT 377

Qy 387 TAGCACAAAGGTATTATTAAGACTTAAATTAATAATGCTCTACGATTGAATGAAAG 446
Db 378 TAGAAATGAAAAGTTTACGATCAATATCTTAAATTCGACC-----AACAG 424

Qy 447 TTTCGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
Db 425 TTCTACTGTACAGTTTTTCAAAATTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 484

Qy 506 GTAAAGAGAGAGGTAAAAACGCTTAGAAGACACAGATAAGAAATTTGAAGATATA 565
Db 485 GGAAAGAAAAGAGAGCTGAAGATTACTTAAAGAGTACGATGATAAGTAGCTGCAATCC 544

Qy 566 AAAAAGA---AATAACTATGATAAAAAATCAAAAGGTATTGCTGCAGTAGCTGCTAAAT 622
Db 545 AAAAGATGCAAAAGCAAGTATAAAGATGATGCGCATTTGAAGCTTCAGTTGTAACT 604

Qy 623 CAGTTTGCCTGCTCATCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
Db 605 TCCGT---GCTGATCATACAGAAATTTATGCTGGTGGATATGCTGGTGAATCTTAAATG 661

Qy 683 TTAAGAGAGCATTAAGTATGATGTTTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTT 742
Db 662 ATTTAGGATTCAAACGTTAATAAGACTTTACAAAAACAAGTTGATTAATCGTAAAGATATA 721

Qy 743 ACTTACAAATGAACACTGAACTTTATCTCAAGTGAATCCTGAGGATGATGTTCAATAA 802
Db 722 TCCAACTTACATCTTAAGAGAGCATTCATTAATGAACGCTGATCATATTTTGTAGTAA 781

Qy 803 CAAACAAAGCAAGTTTCAACGAACCTTCACCTA 834
Db 782 AATCAGATCCAAATGCGAAGATGCTGCATTA 813
```

RESULT 10

US-08-956-171E-238
; Sequence 238, Application US/08956171E
; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 238:
SEQUENCE CHARACTERISTICS:
LENGTH: 3775 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-08-956-171E-238

Query Match 12.3%; Score 124; DB 4; Length 3775;
Best Local Similarity 50.7%; Pred. No. 1.8e-17;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

```
Qy 27 AATTTTAAAGTGAATTCGCTTATTGTTTAAATGCAACTGCAGCATGTGGAAATAA 86
Db 60 AATTTAAATGCTTGTGTTAGCTTGTCTTCTACTTGTTTTAGCAGATGTAGTGGAA 119

Qy 87 TAGTTCAAGTAACTCAAGTAAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
Db 120 TTCAAATAAACAATCATCTGATAACAAGATAAGGAAACAACCTTCAATTAACAATGCAAT 179

Qy 147 AGGTACTACGAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGTAGTATTCAATTGT 206
Db 180 GGGTACAACTGAAATTAAGGGAACCAACGCGTGTGTTACGCTATATCAAGGTGCCAC 239

Qy 207 TGATGCGTTAGTGTCTTTAGATGTTTAAACCTGTTGGGATAGCGGATGATAACAATAA 266
Db 240 TGACGTGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAAATCATCGACACAAA 299

Qy 267 TCGTATTATTAACCATTAAGAGATAAAATTTGGAATAATACACTTCTGTAGGAACACGTA 326
Db 300 ACCGAATTCGAATACATAAAAAATGATTTAAAGATACTAAGATTGTAGTCAAGAAC 359

Qy 327 GCAACCTAACTTAAAGAAATCAGTAACTTAAACCCAGATTTAATTATGTTGCTGATAATA 386
Db 360 TGCACCTAACTTAGAGGAATCTCTAAATTAACCCGAGACTTAAATTTGCGGTCAAAAGT 419

Qy 387 TAGCACAAAGGTATTATTAAGACTTAAATAAATTTGCTCTCCTACGATTGAATGAAAG 446
Db 420 TAGAAATGAAAAGTTTACGATCAATATCTTAAATTCGACC-----AACAG 466

Qy 447 TTTCGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
Db 467 TTCTACTGTACAGTTTTTCAAAATTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 526
```



```
; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3782

Query Match      10.9%; Score 109.4; DB 4; Length 3046;
Best Local Similarity 95.0%; Pred. No. 2.4e-14;
Matches 113; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGAGTGGAAATCAGTGAGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGTTTGTATA 60
Db 2705 GGAGTGGAAATCAGTGAGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGTTTGTATA 2764

Qy 61 ATTGCAACTGAGCATGTGGAAATAATAGTTCAAGTAATCAAGTAAAGAGTCAATCAA 119
Db 2765 ATTGCAACTGAGCATGTGGAAATAATAGTTCAAGTAATCAAGTAAAGAGTCAATCAA 2823

RESULT 13
US-09-543-681A-2690
; Sequence 2690, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIORITY FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2690
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2690

Query Match      7.7%; Score 77.2; DB 4; Length 1056;
Best Local Similarity 48.7%; Pred. No. 1.6e-07;
Matches 270; Conservative 0; Mismatches 278; Indels 6; Gaps 2;

Qy 24 AAAAAATTTAAGTGAATTTGGCTTATTTGTTTAAATTTGCAACTGAGCATGTGGAAA 83
Db 105 AAAAAATTTAAGCCCACTGTTCTTATTTGCTGTTCTCTGTTTATCGCAGGTTGTGTAA 164

Qy 84 TAATAGTTCAAGTAATCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACA 143
Db 165 CGCAAGAATACCTCAACACGCAATCCACAGAGAAACAACCTCTCACTATCGAATGCG 224

Qy 144 AGAAGGTACTACGAAAGTACCTTAAACACCCCTAAACGTTGTTGTTTCTTGGATTTCA 203
Db 225 TCAAGGCACCACTGAGATCCCTGCTCACCACAAAAGTGGTTGTGATGACATGGAAC 284

Qy 204 TGTTGATGCGTTAGTTGCTTTAGATGTTTAAACCTGTTGGATAGCGGATGATAACAAAA 263
Db 285 ACTTGATATTTGTGATGCTCTTGGGCTACCTGTTGTCCTTACCACAAACAAACGTC 344

Qy 264 AAATCGTATTATAACCAATTAAGAGATAAATTTGGAATAATCTCTGTGTAAGAACAG 323
Db 345 CCTACTAAATTTCTAGAGAAATAACCAACGAAATGAGTACATCAACGAAGGTGGCT 404

Qy 324 TAAGCAACCTACTTTAGAGAAATCAGTAAACTTAAACCCAGATTTAAATTTATGCTGATA 383
Db 405 GTTTGAACCAAACTACGAAAACTCAGTACTACTGCTCCCTGATCTGATTTTAAACGGTAG 464

Qy 384 TAATAGACACAAGGATTTTATAAGACTTAAATAAATTTGCTCTCTCAAGTTGAAGTGA 443
Db 465 TGTGCTCGC---GATGATATGCGCAAAATTAAGTGAATTCGACCATCCATTTCTATGGA 521

Qy 444 AAGTTTCGATGAGATTAATTAAGAAAAATTTGATGCTTTTAAAAACAATTTCA---AAGC 500
```

```
Db 522 TATCGACAGCACACGCTTTATTGTGACAGCTCACTGAGCGTACAAAGCACTTTAGGTCAAAT 581
Qy 501 TTTAGGTAAAGAAAGAGGTAAAAAACGGTTTAAAGAACACACGATTAAGAAAAATTGAAGA 560
Db 582 TTTTGGTAAAGAGAAACAAAGCTAAAAAAATTTATTAGCTGACTTCAATAGCAAAAATTGTATAC 641
Qy 561 ATATAAAAAAGAAA 574
Db 642 GGTAAAAAGCAAAA 655

RESULT 14
US-08-956-171E-383/C
; Sequence 383, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 383:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 383:
US-08-956-171E-383

Query Match      7.3%; Score 73.8; DB 4; Length 421;
Best Local Similarity 61.9%; Pred. No. 7.1e-07;
Matches 117; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 7 GAATCAGTGAGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGTTTAAATTGCA 66
Db 205 GGAAGATGAGAGGTTCTAAAACTTTTAGTATATTGGGATTAATAGTTGCGCTTACTTTA 146
Qy 67 ACTGAGCATGTGGAAATAAATAGTTCAAGTAACTCAAGTAAAGATCATCAAAAGATGA 126
Db 145 GTTCGAGCTTGTGGTAAATACGGATAATCAAGTAAAAAAGAAATCATCAACTAAGATACT 86
```

QY 127 GTTGAATCAACGACGAAAGAGTACTACGAAAGTACCTAAACACCCCTAAACGTTGTTT 186
Db 85 ATTTTCGGTAAAGAGTGAATAATGTTACAGTAAAGTACCTAAAGATGCAAAACGTTATCGTT 26
QY 187 GTTCTTGAG 195
Db 25 GTATTAGAG 17

RESULT 15
US-08-781-986A-383/c
; Sequence 383, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesh
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 383:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-383

Query Match 73.3%; Score 73.8; DB 4; Length 421;
Best Local Similarity 61.9%; Pred. No. 7.1e-07;
Matches 117; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 7 GAATCAGTCAGAGGTTTAAATAATTTAAAGTAAATGCGCTTATGTTGTTTAAATGCA 66
Db 205 GGAACGATGAGAGGCTCTAAACCTTTTAGTATATGCGGATTATAGTTCCTTACTTTTA 146
QY 67 ACTGCAGCATGTGGAATAATAGTTCAAGTAACCTCAAGTAAAGAGTCATCAAAAGATGGA 126
Db 145 GTTGCAGCTGTGTTAATACGGATATTCAAGTAAAGAAATCATCAACTAAAGATACT 86
QY 127 GTTGAATCAACGACGAAAGAGTACTACGAAAGTACCTAAACACCCCTAAACGTTGTTT 186
Db 85 ATTTTCGGTAAAGAGTGAATAATGTTACAGTAAAGTACCTAAAGATGCAAAACGTTATCGTT 26
QY 187 GTTCTTGAG 195
Db 25 GTATTAGAG 17

RESULT 16

US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
; US-09-902-540-1357

Query Match 7.2%; Score 72.2; DB 4; Length 612;
Best Local Similarity 50.4%; Pred. No. 1.7e-06;
Matches 173; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 254 ATACCAAAAAAATCGTATTATTAAACCATTTAAGAGATAAAATTTGGAATAACACTTCTG 313
Db 150 ATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 209
QY 314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAACCTTAAACACAGATTTAATTA 373
Db 210 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 269
QY 374 TTGCTGATAATAATAGACACAAAGGTATTTATAAGACTTTAAATAAAAATTTGCTCTACGA 433
Db 270 ATAAAAAATAAAAAATTTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 329
QY 434 TTGAACCTGAAAGTTTCGATGAGATTATATGAATAATTTGATGCTTTTAAAAACAATTT 493
Db 330 AAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 389
QY 494 CAAAGCTTTAGGTAAAGAAAGAGGCTTAAAGAAACGCTTAGAAGAAACACGATAAGAAAA 553
Db 390 ATAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 449
QY 554 TTGAAGAAATATAAAAAAGAAATAACTATGATAAAAAATCAAAA 596
Db 450 ATTAATAATAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 492

RESULT 17
US-09-071-035-37
; Sequence 37, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2

```
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/071.035
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: A. Anders Brookes
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB369P2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 37:
;; SEQUENCE CHARACTERISTICS:
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-071-035-37

Query Match      7.0%; Score 70.6; DB 3; Length 975;
Best Local Similarity 48.8%; Pred. No. 4.1e-06;
Matches 317; Conservative 0; Mismatches 314; Indels 18; Gaps 4;

Qy 48 ATTGTTGTTTAAATTCGCACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 ACTGTTAATGTTAAGTGTCTGCAACAAATAAAAAAACACGACGATTCTGCAACACAGA 125
Qy 108 AGAGTCATCAAAAGATGGAGTTGAATCAACGACGAGAGGACTACTAGAAAGTACCTAA 167
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126 AACCAACAGCTAAAAACGGAAGTCACAGTCAAAAGACACCAATGGTCAATTAACCGTCCCAA 185
Qy 168 ACACCTTAAACGTTGTTGTTCTTGAGTATTCATTTGTCATGGTTAGTTGCTTTAGA 227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
186 AATCTCTAAGAAAGTCGTGTTGTTTGTGATATGTTGTCGATACATGGATGCACTAGG 245
Qy 228 TGTGTTAAACCTGTTGGGATAGCGGATGATTAACAAAAAATCGTATTAATAACCAATTAG 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
246 TGTCCGTGACCGCGTGGTAGTGCGCAACTAAAAATATCCCTGGTATTTGAAAAAATA 305
Qy 288 AGATATAAATTGGAAAAATACACTCTGTAGGAAACAGCTAGGCAACTACTACTTAGAGAAAT 347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 CCAAAAAGTTGAA-----TCAGCAGCGCGCATTAAGAACCAGATTTAGAAAAAAT 356
Qy 348 CAGTAAACTTAAACAGATTTAATTTATGCTGATAATAATAGACACAAAGGTATTATAA 407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
357 CAATCAACTAAACAGACTTAATTAATTTCTGTCGTCAC-----AGATTATCAAGA 413
Qy 408 AGACTTAAATAAATTCCTCTACGATTGAATGAAATGAAAGTTTCGATGGAGATTATAATGA 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 ACAATTAAGGCCATTGGCGCAACCAATTACTTAGCTAGATGCCAAAAATCCTTGGGC 473
Qy 468 ---AAATATTGATGCTTTTAAAACAATTTCAAAGCTTTTAGTTAAAGAGAGAGAGGTAA 524
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
474 ATCAACGAAACAAAATATCGAAACGTTAGGCACATAATTTTATAAAGAAAGAGGTAGCTAA 533
Qy 525 AAAACGCTTAGAAGAACACGATAAGAAATTTGAAGAATATAAAAAAGAAATAAATCTATGA 584
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
534 AGAAAAATAACTGGCTTAGAAAAAGAAATTTGCTGACCTGAAAAAACA---AGCAGAGC 590
Qy 585 TAAAAATCAAAAGGATTGCTCGAGTAGCTGCTAAATCAGGTTTGTGCTCATCCAAAG 644
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
591 TAGCGGAATAATGCGCTTGTGTTAGTTTAAACGAAGGACAACTTCCGCTTACGGAAA 650
Qy 645 CAACCTTATGTTGGTCAATTCCTTAAGTCAACTAGGTTTAAAGAGCA 693
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
651 AGGCTCTCGTTTCGGTTTAAATTCATGATACATTTGGCTTCAAGAGCAGCA 699
```

RESULT 18

```
US-09-134-000C-937
; Sequence 937, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 937
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-937
```

```
Query Match      7.0%; Score 70.6; DB 4; Length 984;
Best Local Similarity 48.8%; Pred. No. 4.1e-06;
Matches 317; Conservative 0; Mismatches 314; Indels 18; Gaps 4;

Qy 48 ATTGTTGTTTAAATTCGCACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 ACTGTTAATGTTAAGTGTCTGCAACAAATAAAAAAACACGACGATTCTGCAACACAGA 122
Qy 108 AGAGTCATCAAAAGATGGAGTTGAATCAAGCAGAGAGGTTACTAGAAAGTACCTAA 167
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 AACCAACAGCTAAAAACGGAAGTCACAGTCAAAAGACACCAATGGTCAATTAACCGTCCCAA 182
Qy 168 ACACCTTAAACGTTGTTGTTCTTGAGTATTCATTTGTCATGGTTAGTTGCTTTAGA 227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 AATCTCTAAGAAAGTCGTGTTGTTTGTGATAATGTTCTTGGATACAAATGGATGCACTAGG 242
Qy 228 TGTGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAATCGTATTAATAACCAATTAG 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 TGTCCGTGACCGCGTGGTAGTGCGCAACTAAAAATATCCCTGGTATTTGAAAAAATA 302
Qy 288 AGATAAATTTGGAAAAATACACTTCTGTAGGAAACAGCTAAGCAACTAATCTTAGAAGAAAT 347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 CCAAAAAGTTGAA-----TCAGCAGCGCGCATTAAGAACCAGATTTAGAAAAAAT 353
Qy 348 CAGTAAACTTAAACAGATTTAATTTATGCTGATAATAATAGACACAAAGGTATTATAA 407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
354 CAATCAACTAAACCCAGACTTAATTAATTTCTGTCGTCAC---AAGATTATCAAGA 410
Qy 408 AGACTTAAATAAATTTGCTCTACGATTGAACGAAAGTTTCGATGGAGATTATAATGA 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
411 ACAATTAAGGCCATTGGCGCAACCAATTACTTAGCTGTAGATGCCAAAAATCCTTGGGC 470
Qy 468 ---AAATATTGATGCTTTTAAAACAATTTCAAAGCTTTTAGGTAAAGAAAGAGGTAA 524
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 ATCAACGAAACAAAATATCGAAACGTTAGGCACATAATTTTGTATAAAGAAAGAGGTAGCTAA 530
Qy 525 AAAACGCTTAGAAGAACACGATAAGAAATTTGAAGAATATAAAAAAGAAATAAATCTATGA 584
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
531 AGAAAAATAACTCGCTTTAGAAAAAGAAATTTGTCGCTGAAAGAAACA---AGCAGAGC 587
Qy 585 TAAAAATCAAAAGGTATTGCTCGCAGTAGCTGCTTAAATCAGGTTTGTGCTCATCCAAAG 644
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
588 TAGCGGAAATAATGCGCTTGTGTTAGTTAGTTAAACGAGGACAACTTCCGCTTACGGAAA 647
Qy 645 CAACCTTATGTTGGTCAATTCCTTAAGTCAACTAGGTTTAAAGAGCA 693
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
648 AGGCTCTCGTTTCGGTTTAAATTCATGATACATTTGGCTTCAAGAGCAGCA 696
```

```
RESULT 19
US-09-902-540-1583
; Sequence 1583, Application US/09902540
; Patent No. 6833447
```

GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1583
; LENGTH: 709
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1583

Query Match 6.9%; Score 69.8; DB 4; Length 709;
Best Local Similarity 49.9%; Pred. No. 5.7e-06;
Matches 210; Conservative 0; Mismatches 202; Indels 9; Gaps 1;

QY 264 AATCGTATTATTAACCACTTAAGAGATAAAATGGAATAATACACTTCTGTAGGAACAG 323
DB 7 AATCCGCGGTTTGACCATATTAAAGGTCAGATGAAGCGGTTAAAGAGTGTGGGTACAGA 66
QY 324 TAAGCAACTTACTTAGAAGAAATCAGTAACTTAACCCAGATTTAAATATTGCTGTATAA 383
DB 67 AAGCGAGCCCAAGCTTAGAAGCAATCGCTGCTTTAAAGCGCTGACTTGTATTCGGAATAA 126
QY 384 TAATAGACACAAGGATTATTAAGAGACTTAATAAATGCTCTCAGATTGAACTGAA 443
DB 127 AATGGACAAAGAAAGATATAGATCAGTCAGCTAGTCAATTTGCTCCGACTGATTTCTGA 186
QY 444 AAGTTTGCATGAGATTAATAAGAAATATGATGCTTTTAAACAAATTTCAAAGCTTT 503
DB 187 GCGAGTGGCGGAGATTGGAATCAAC-----TTTAAGCTATATGCAAAAGCTGT 237
QY 504 AGGTAAGAGAAAGAGGTTAAAGACGCTTAGAAGACACAGATGAAGAAATGAAGATA 563
DB 238 AAACAAGAAAGAAAGGTTAAAGAGATATTAGCTGACTATGCAATCGGTAGCAGATT 297
QY 564 TAAAAAGAAATAAATATGATGATAAAATCAAAAGGATTGCTCGAGTAGCTGCTAAATC 623
DB 298 AAGAAACCCCTTGGAGATCAGTTAAACAAAATACTCAGTTGACGCTTTACAGCTGG 357
QY 624 AGGTTTGCCTGCTATCCCAAGCAACTTTATGTTGCTCAATTCCTAAGTCAACTAGGTTT 683
DB 358 AGAGGTTGCTATCTATCATATAAAGATTCTCTCAGGTGTTATTTAGATCAGCTTGGATT 417
QY 684 T 684
DB 418 T 418

RESULT 20
US-09-902-540-2208
; Sequence 2208, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2208
; LENGTH: 702

; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-2208

Query Match 6.8%; Score 68.8; DB 4; Length 702;
Best Local Similarity 49.8%; Pred. No. 9.4e-06;
Matches 209; Conservative 0; Mismatches 202; Indels 9; Gaps 1;

QY 265 AATCGTATTATTAACCACTTAAGAGATAAAATTCGAAAAATACACTTCTGTAGGAACAGT 324
DB 1 AATCCGCGGTTTGACCATATTAAAGGTCAGATGAAGCGGTTAAAGAGTGTGGGTACAGA 60
QY 325 AAGCAACTAATCTTAGAAGAAATCAGTAACTTAACCCAGATTTAAATATTGCTGTGATAT 384
DB 61 AGCGAGCCCAAGCTTAGAAGCAATCGCTGCTTTAAAGCGCTGACTTTGATTATCGGAATAA 120
QY 385 AATAGACACAAGGATTATTAAGAGACTTAATAAATTTGCTCTACGATTGAACGTAAA 444
DB 121 ATGCGACAAAGAAAGATATAGATCAGCTTAGTCAATTTGCTCCGACTGTATTCTCTGAG 180
QY 445 AGTTTTCGATGAGATTTAATAAGAAATATTGATGCTTTTAAAAACAATTTCAAAGCTTTTA 504
DB 181 GAGTGGCGGAGATTGGAATCAAC-----TTTAAGCTATATGCAAAAGCTGTA 231
QY 505 GGTAAAGAGAAAGAGGTAAAAACGGTTAGAGAAACACAGATAAGAAATTTGAAGAAATAT 564
DB 232 AACAAGAGAAAGAGGTAAAGAGATATTAGCTGACTATGACAAATCGCGGTAGCAGATTTA 291
QY 565 AAAAAAGAAATACTATGGATAAAANTCAAAAGGTATTGCTCGAGTAGCTGCTAAATCA 624
DB 292 AAGAAACCCCTTGGAGATCAGTTAAACAAAATAATCTCAGTTGTACGCTTTTACAGCTGA 351
QY 625 GGTTCGTTGCTCATCCAAAGCAACTCTTATGTTGCTCAATTCCTAAGTCAACTAGGTTT 684
DB 352 GACGTTGCTATCTATCATAAAGATTCATTCAGGTGTTATTTAGATCAGCTTGGATT 411

RESULT 21
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:


```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3348
; LENGTH: 3153
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3348

Query Match          6.4%; Score 64.6; DB 4; Length 3153;
Best Local Similarity 47.0%; Pred. No. 0.0001;
Matches 452; Conservative 0; Mismatches 479; Indels 30; Gaps 7;

QY 61 ATTGCAACTGCAGCATGTGGAATATAGTTCAGTAACCTCAAGTAAGAGTCATCAAAA 120
DB 1188 AGTGATCTTAAAGAACTGTACCATCAAAATAGTTTGAAGCAAGTGTGAAGAAAT 1129

QY 121 GATGGAGTTGAAATCAAGCAGCAAGAGGTACTACGAAAGTAGTACCTAAACACCCCTAAACGT 180
DB 1128 AATGGCAGTGATGAAGAAAATCTCTAATCTGTCGAAGTACCAAGAAATCCTAAAT 1069

QY 181 GTTGTGTTCTTGAGTAATTCATTTGTGTGAGCGTTAGTTGCTTTAGATGT-----TAAA 234
DB 1068 GCCGTTGTATTAGATTATCGAGCGCTTGATGTGTTGAAAGAAATTAGGTGTGGCTGATAAA 1009

QY 235 CCTGTTGGATAGCGGATGATACAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
DB 1008 GTAAAAGGTTTACCTTAAGAGTGAATAACCAATCTTTACCTAAATTTTAGATGAATTT 949

QY 292 AAAATTGGAATAATACACTTCTGTAGGAACACGCTAAGCAACCTTAACCTTAGAAGAAATCACT 351
DB 948 AAGATGATGAATGATATTAACTCGGAATTTAAAGAAAGTGAACCTTGTGATAAGTTGCA 889

QY 352 AAACTTAAACCAAGATTAA---TTATTGCTGATAATTAATAGACACAAAGGTATTATATAA 408
DB 888 TCAGCTAAACCAAGATGTATTTTATTTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 829

QY 409 GACTTAAATAAATTGCTCTAGTGTGAACTGAAAGTTTCGATCGAGATTATATAGAA 468
DB 828 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGTCAAGATGATGACAAC 769

QY 469 AATATTGATGCTTTTAAACAAATTTCAAAAGCTTTTAGGTAAAGAAAGAGTGAATAAAA 528
DB 768 TTAATTAAAGATATGAATAAATAATACAGAAATTTAGGGAATCTACGA---TAAAGAA 712

QY 529 CGCTTAGAGAAACACGATAGAAAATTTGAAGAAATATAAAGAAATAACTATGGATAAA 588
DB 711 GATAAAGCTAAAAAATTAATAAAGATTTAGATAGAAAATACTCGATATGAAGATAAA 652

QY 589 AATCAAAAGGTATGCTCGAGTAGCTGCTAAATCAGGTTTGTCTCATCAAGCAAC 648
DB 651 ACTAAAAGACTTTAATAAGAAAGTAATGTATTTATTGTTTAAAGAGGTGAACCTATCAACG 592

QY 649 TCTTATGTTGCTCAATTCCTAAGTCAACTAGGTTTAAAGAACATTTAAGTGTATGATTT 708
DB 591 TTTGGACCGAGGAAGATTTGGTGGTTTGTAGTGTGTGTATACATAGGATTAACCTGCA 532

QY 709 ACTAAAAGGTTTAAAGTAAGTATCTTTAAAGGACCTTACTTACAAATGAACACTGAACTTTA 768
DB 531 GACAAAAGGTTAGCAAAAGCCCGCATGGTC-----AAAATATAAATGAATATATTT 478

QY 769 TCTCAAGTGAATCTCTGAGCGGTATGTTTCAATATGACAAACAAAGCAAGTCTTCAACG---AA 825
DB 477 AACCAAGCAGAAATCCAGATGTATTTTAGCTATGATGTTGTATACATAGGATTAACCTGTA 418

QY 826 CCTTCACTAAAGAACTAGAAAGATCCCTGTATGGAAGAAATTTAAACGCTGTGGAATAAT 885
DB 417 GCAACAACAATCAAGTTTAAAAAACAAGTTTATAAAAAATGTAAAGAAAGTAAAGAT 358

QY 886 CAACGCTGTGATATTTTAAAGCCGTGACTTATGGCAAGATCACTGTGTTTAAATTTCTTCA 945
DB 357 AATCATATTTACGAATTAGATCCAAACTATGTTTCTC-----TTCAGGATCTTCA 304
```

```
QY 946 GAAGAAATGCAAAAGAACTTTGTTGAATATTCTAAGAAAGATAGTAAAAAGATAATAAG 1005
DB 303 ACGACAACTATCAAAACAAATTTGATGAATTAATGAAGTAGTAGAGAAAGTTGAAAAATAA 244

QY 1006 T 1006
DB 243 T 243

RESULT 24
US-09-710-279-4042
; Sequence 4042, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4042
; LENGTH: 3267
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4042
```

```
Query Match          6.4%; Score 64.6; DB 4; Length 3267;
Best Local Similarity 47.0%; Pred. No. 0.0001;
Matches 452; Conservative 0; Mismatches 479; Indels 30; Gaps 7;

QY 61 ATTGCAACTGCAGCATGTGGAATATAGTTCAGTAACCTCAAGTAAGAGTCATCAAAA 120
DB 800 AGTGATCTTAAAGAACTGTACCATCAAAATAGTTTGAAGCAAGTGGTAAAGAAAT 859

QY 121 GATGGAGTTGAAATCAAGCAGCAAGAGGTACTACGAAAGTAGTACCTAAACACCCCTAAACGT 180
DB 860 AATGCAAGTGAAGAAAAAATCTCTAATACTGTGGAAGTACCAAGAAATCCTAAAAAT 919

QY 181 GTTGTGTTCTTGAGTATTTCATTTGTCATGCTGTTAGTTGCTTTAGATGT-----TAAA 234
DB 920 GCCGTTGTATTAGATTATGAGCGCTTGAATGTTTGAAGAAATTTAGGTGTGCTGTATAA 979

QY 235 CCTGTTGGATAGCGGATGATAACAAAAAAATCGTATTATTATTAACCATTAAGAGA---T 291
DB 980 GTAAAAGGTTTACCTAAGGTGAATAAACCAATCTTTACCTAAATTTTAGATGAATTT 1039

QY 292 AAAATTGGAATAATACACTTCTGTAGAACACGCTAAGCAACCTTAACCTTAGAAGAAATCACT 351
DB 1040 AAGATGATAGTATATTAATTAATCTGGAATTTTAAAGAAAGTGAACCTTTGATAAGTTGA 1099

QY 352 AAACCTTAACCAAGATTAA---TTATTGCTGATAATATAGACACAAAGGTATTATATAAA 408
DB 1100 TCAGCTAAACCAAGATGTGATTTTATTTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 1159

QY 409 GACTTAAATAAATTTGCTCTTACGATTGAACTGAAAGGTTTCGATGGAGATTATATAGAA 468
DB 1160 GAATTTAAAGAAAGTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAGTGTAGCAAC 1219

QY 469 AATATTGATGCTTTTAAAAACAATTTCAAAAGCTTTAGGTAAAGAAAGAGGTAAGAA 528
DB 1220 TTAATTAAAGATATGAAAAAATAACAGAAAAATTTAGGCAAAATCTACGA---TAAAGAA 1276

QY 529 CGCTTAGAAGAACACGATGAAGAAATTTGAAGAAATTAAGAAATATAAAGAAATTAATCTAGTAA 588
DB 1277 GATAAAGCTAAAAAAATTAATGAAGATTTTAGATAGAAAAATATCTGTATGAAGATAAA 1336
```

589 AATCAAGAGTATTCCTGCGAGTAGCTGCTAAATCAGGTTTGGCTTCATCCCAAGCAAC 648
1337 ACTAAGAGCTTTAATAAGAAAGTAATGTAATTTATTTGGTTAAAGAGGTGAAGCTATCAACG 1396
649 TCTTATGTTGCTCAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTAAGTCATGATGTT 708
1397 TTGGACCAAGGAGGAGATTTGGTGGTTTAAAGAGTATTAAGGATTTAAACCTGCA 1456
709 ACTAAGAGTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAAACACTGAACTTTA 768
1457 GACAAAAGGTTAGCAAAAGCCGATGCTC-----AAAATATAAATAATGAATATATT 1510
769 TCTCAAGTGAATCCTGAGCGTATGTTTCAATATGATGACAAACAAAGCAAGTTCACG--AA 825
1511 AACAAAGCAGATCCAGATGTTTATTTAGCTATGATCGTGTTCAGTTGTTAGGTGGTAAA 1570
826 CTTTCACTAAAGAACTAGAAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAAT 885
1571 GCAACACAAATCAAGTTTAAAGAAACAAAGTTATAAAGATGTAAAGAGCAGTAAAGT 1630
886 CAACGTGTTGATATTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAAATTTCTTCA 945
1631 AATCATATTTACGAATTAGATCCAAACTATGTTTCTC-----TTCAGGATCTTCA 1684
946 GAAGAAATGCGCAAGAACTGTTGAATATCTAAGAAAGATAGTAAAGAAAGATATAAG 1005
1685 ACGAACTATCAACAAATGATGAATTAATGAAGTAGTAGAGAAAGTTGAAAAATAA 1744
1006 T 1006
1745 T 1745

RESULT 25

US-09-710-279-3564
; Sequence 3564, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3564
; LENGTH: 3618
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3564

Query Match 6.4%; Score 64.6; DB 4; Length 3618;
Best Local Similarity 47.0%; Pred. No. 0.00011;
Matches 452; Conservative 0; Mismatches 479; Indels 30; Gaps 7;
61 ATTGCACTGCGAGCATGTGGAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
666 AGTGATTTAAAGAACTGTAAACCATCAAAATAGTTTGAAGCAAGTGGTAAAGAAAT 725
121 GATGAGTTGAAATCAAGCACGAGGAAGGTACTAGCAAGTACCTTAAAGCAACCCCTAAAGT 180
726 AATGCACTGATAGCAAAAATCTTAATCTGTCGAAGTACCAAGAAATCTTAAAT 785
181 GTTGTGTTCTTGAATATTCATTTGTTGATGGGTTAGTGTCTTTAGATGT-----TAAA 234
786 GCGGTTGTATAGATTATGAGCGCTTGTATGTGTTGAAAGAAATTAGGTGTGGCTGATAAA 845
235 CCGTGTGGATAGCGGATGATGAACAAAAAAATCGTATTATTAAACCATTAAGACA--T 291

846 GTAAAAGGTTTACCTAAGGTGAAATTAACCAATCTTTTACCTAAATTTTATAGATGATTT 905
292 AAAATTCGAAATAACATCTTCTGTAGGAACAGTGAAGCAACCTTAACCTTAGAAGAATTCAGT 351
906 AAGATGATAAGTATATAATTAATCTGGAATTTTAAAGAAAGTGAACCTTTGTATAAAGTTGCA 965
352 AAACCTTAAACCAAGATTTAA--TTATTTGCTGATTAATATAGACACAAAGGTATTTATAAA 408
966 TCAGCTAAACCAAGATGATTTTATTTTCAGGAAGAAACAGCTAATCAGAAAAATTTAGAT 1025
409 GACTTAAATAAAATTCCTCCTACGATTTGAATCTGAAAGTTTTCGATGGAGATTTAATGAA 468
1026 GAATTTAAAGAGCTGACCAAAAGCTAAAGTTGATATATGTAGGTACAAAGTGTGACAAAC 1085
469 AATATTCATGCTTTTAAACAAATTTCAAAAGCTTTTAGTAAAGAAAGAAAGGTAAAGAAA 528
1086 TTAATTAAGATATGAAAAAATAACAGAAAAATTTAGGAAAAATCTACGA---TAAAGAA 1142
529 CGCTTAGAAGNACAGATTAAGAAATTTGAAGATATTAAGAAAGAAATTAACCTATCGATAAA 588
1143 GATAAGCTAAAAAATTAATAAGATTTTAGATAGAAAAATATCTGATATGAAAGATATAA 1202
589 AATCAAAAGGTATTGCTGCAAGTAGCTGCTAAATCAGGTTTGTCTTCATCCCAAGCAAC 648
1203 ACTAAGACTTTAATAAGAAAGTAAATGTTATTTATTTGTTAAAGAGGTGAACATCAACG 1262
649 TCTTATGTTGCTCAATTCCTAAGTCAACTAGGTTTAAAGAAAGCAATTAAGTGTGATGTT 708
1263 TTTGGACCAAGGAGGAGATTTGGTGGTTTAGTGTGATACATTTAGGATTTTAAACCTGCA 1322
709 ACTAAGGTTTAACTAAGTATCTTAAAGGACCTTACTTACAAATGAACACCTGAACCTTTA 768
1323 GACAAAAGGTGACAAAAGCCGCGATGGTC-----AAAATATAAATAATGAATATATT 1376
769 TCTCAAGTGAATCCTGAGCGTATGTTCAATATGATCAAAACAAAGCAAGTTCTTAAAG--AA 825
1377 AACAAAGCAGATCCAGATGTTATTTAGCTATGATCGTGGTTCAGTTGTAGGTGGTAAA 1436
826 CTTTCACTAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAGCGCTGTGAAAAAT 885
1437 GCAACAACTAATCAAGTTTAAAGAAACAAAGTTTAAAGAAATGTAAGAAAGCAGTAAAAAGT 1496
886 CAACGTGTTGATATTTTAGACCGTGACTTATGCGCAAGATCACGTGGTTTAAATTTCTTCA 945
1497 AATCATATTTACGAATTAGATCCAAACTATGGTATTCTC-----TTCAGGATCTTCA 1550
946 GAAGAAATGCGCAAAAGAACTTGTTCATTTAAGAAAGATAGTAAAAAGAAAGATAAAG 1005
1551 ACGCAACTATCAACAAATTTGATGAATTAATGAAGTAGTAGAGAAAGTTGAAAAATAA 1610
1006 T 1006
1611 T 1611

RESULT 26

US-09-710-279-461
; Sequence 461, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 461
; LENGTH: 1044
; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-09-710-279-461

Query Match 6.4%; Score 64.4; DB 4; Length 1044;

Best Local Similarity 47.1%; Pred. No. 9e-05;

Matches 450; Conservative 0; Mismatches 476; Indels 30; Gaps 7;

```
QY 61 ATTGCAACTGCGAGCTGTGGAATAATAGTTCAGTAACCTCAAGTAAGAGTCATCAAAA 120
DB 100 AGTGATTCTAAAGAACTGTAAACCATCAAAAATAGTTTGAAGCAAGTGTGAAGAAAT 159
QY 121 GATGGAGTTGAATCAAGCAGCAGAGGCTACTACGAAAGTACCTAAACACCCCTAAAGCT 180
DB 160 AATGGCAGTGATAAGAAAAAATCTCTAATACTGTGCAAGTACCAAGAAATCCTAAAT 219
QY 181 GTTGTGTTGTTCTTGAGTATTCAITTTGTTGATGGCTTAGTTGCTTTAGATGT-----TAAA 234
DB 220 GCGGTGTATTAGATTATGGAGCGCTTGATGTGTGAAGAAATTAGTGTGGCTGATAAA 279
QY 235 CTTGTTGGGATACGGATGATACAAAAAATTCGTATTATTAAACCATTAAGAGA---T 291
DB 280 GTAAAAAGGTTTACCTAAAGGTGAAAAATAACCAATCTTTACCTAAAAATTTTAGATGAATTT 339
QY 292 AAAATTGGAAATACACTTCTGTAGGAACAGCTTAAGCAACCTTAACCTTAGAAGAAATCACT 351
DB 340 AAAGATGATAAGTATTAATACTGGAATTTTAAAGAAAGTGAACCTTTGATAAGTTGCA 399
QY 352 AAATCTTAAACCAAGTTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTATTATAAA 408
DB 400 TCAGCTTAAACCAAGTGTATTATTTTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 459
QY 409 GACTTAAATAAATTTCTCTACGATTCGAACCTGAAAGTTTCGATGAGATTAATAGAA 468
DB 460 GAAATTTAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAAGTATCAAC 519
QY 469 AATATTGATGCTTTTAAACCAATTTCAAAGCTTTTAGGTAAAGAAAGAAAGTAAAGAA 528
DB 520 TTAAATTAAGATATGAAAAAATTAACAGAAATTTTAGGGAAATCTACGA---TAAAGAA 576
QY 529 CGCTTAGAAGAACACGATAAGAAAAATTTGAAGAAATATAAAAAAGAAATACTATGGATAAA 588
DB 577 GATAAAGCTTAAAAAATTAATAAGATTTAGATAGAAAAATATCTGATATGAAGATAAA 636
QY 589 AATCAAAAGGTATGCTGCGAGTACGTGCTGCTAAATCAGGTTTTCGCTCATCAAGCAAC 648
DB 637 ACTAAAGACTTTTAATAAGAAAGTAAATGTATTTATTGTTTAAACCAAGGTGAACATCAACG 696
QY 649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTGTTTAAAGAAAGCATTAAGTGTATGTT 708
DB 697 TTTGGACAGGAGGAAGATTGGTGGTTTATGTTGTTGATACATTTAGGATTTAAACCTGCA 756
QY 709 ACTAAAGTTTAAAGTATCTTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
DB 757 GACAAAAAGGTTAGCAAAAGCCGATGGTC-----AAATATAATAATGATGATATATT 810
QY 769 TCTCAAGTGAATCCTGAGCGGTATGTTTCATATATGACAAACAAAGCAAGTTCTAAG---AA 825
DB 811 AACCAAGCAGAAATCCAGATGTTATTTTAGCTATGATCGTGGTTCAGTTGAGTGGTAAA 870
QY 826 CTTTCACTAAAGAACTAGAAAAAGATCTCTGATGGAGAAATTAACCGCTGTGAAGAAAT 885
DB 871 GCAACAAACAAATCAAGTTTAAAAAACAAGTTTATAAAAAATGTAAAGCAGTAAAGAAAGT 930
QY 886 CAACGCTGTGATATTTTAGACCGGTACTTATGGCAAGATCACGTGGTTTAAATTTCTTCA 945
DB 931 AATCATATTTACGAATTTAGATCCAAAACCTATGTTATTTCTC-----TTGAGATCTTCA 984
QY 946 GAAGAAATGGCAAAAGAACTTTGTAATTTATTAAGAAAGATAGTAAAAAAGATAA 1001
DB 985 ACGACAACTATCAAAACAAATTTGATGAATTAATGAAGTAGTAGAAGAAAGTTGAAAA 1040
```

RESULT 27

US-09-710-279-1267

; Sequence 1267, Application US/09710279

; Patent No. 6703492

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: P03480US

; CURRENT APPLICATION NUMBER: US/09/710,279

; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1267

; LENGTH: 1044

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-09-710-279-1267

Query Match

Best Local Similarity 6.4%; Score 64.4; DB 4; Length 1044;

Matches 450; Conservative 0; Mismatches 476; Indels 30; Gaps 7;

```
QY 61 ATTGCAACTGCGAGCTGTGGAATAATAGTTCAGTAACCTCAAGTAAGAGTCATCAAAA 120
DB 100 AGTGATTCTAAAGAACTGTAAACCATCAAAAATAGTTTGAAGCAAGTGTGAAGAAAT 159
QY 121 GATGGAGTTGAATCAAGCAGCAGAGGCTACTACGAAAGTACCTAAACACCCCTAAAGCT 180
DB 160 AATGGCAGTGATAAGAAAAAATCTCTAATACTGTGCAAGTACCAAGAAATCCTAAAT 219
QY 181 GTTGTGTTGTTCTTGAGTATTCAITTTGTTGATGGCTTAGTTGCTTTAGATGT-----TAAA 234
DB 220 GCGGTGTATTAGATTATGGAGCGCTTGATGTGTGAAGAAATTAGTGTGGCTGATAAA 279
QY 235 CTTGTTGGGATACGGATGATACAAAAAATTCGTATTATTAAACCATTAAGAGA---T 291
DB 280 GTAAAAAGGTTTACCTAAAGGTGAAAAATAACCAATCTTTACCTAAAAATTTTAGATGAATTT 339
QY 292 AAAATTGGAAATACACTTCTGTAGGAACAGCTTAAGCAACCTTAACCTTAGAAGAAATCACT 351
DB 340 AAAGATGATAAGTATTAATACTGGAATTTTAAAGAAAGTGAACCTTTGATAAGTTGCA 399
QY 352 AAATCTTAAACCAAGTTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTATTATAAA 408
DB 400 TCAGCTTAAACCAAGTGTATTATTTTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 459
QY 409 GACTTAAATAAATTTCTCTACGATTCGAACCTGAAAGTTTCGATGAGATTAATAGAA 468
DB 460 GAAATTTAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAAGTATCAAC 519
QY 469 AATATTGATGCTTTTAAACCAATTTCAAAGCTTTTAGGTAAAGAAAGAAAGTAAAGAA 528
DB 520 TTAAATTAAGATATGAAAAAATTAACAGAAATTTTAGGGAAATCTACGA---TAAAGAA 576
QY 529 CGCTTAGAAGAACACGATAAGAAAAATTTGAAGAAATATAAAAAAGAAATACTATGGATAAA 588
DB 577 GATAAAGCTTAAAAAATTAATAAGATTTAGATAGAAAAATATCTGATATGAAGATAAA 636
QY 589 AATCAAAAGGTATGCTGCGAGTACGTGCTGCTAAATCAGGTTTTCGCTCATCAAGCAAC 648
DB 637 ACTAAAGACTTTTAATAAGAAAGTAAATGTATTTATTGTTTAAACCAAGGTGAACATCAACG 696
QY 649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTGTTTAAAGAAAGCATTAAGTGTATGTT 708
DB 697 TTTGGACAGGAGGAAGATTGGTGGTTTATGTTGTTGATACATTTAGGATTTAAACCTGCA 756
```


STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLYING APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 9834 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-37

Query Match 6.1%; Score 61.8; DB 4; Length 9834;
Best Local Similarity 47.0%; Pred. No. 0.00052;
Matches 334; Conservative 0; Mismatches 362; Indels 15; Gaps 4;

QY 291 TAAATTTGGAATACACCTCTCTAGGAAACACGTAAGCAACCTAACTTAGAAGAAATCAG 350
DB 28 TAAAGATGATTAATATACAAAGTTGTTAATTTAAAGAGTGAATTTTGATATAATTCG 87
QY 351 TAAACTTAAACACGATTTAA---TTATGCTGATTAATATAGACAAAGGATTTATTA 407
DB 88 TCGACGAAACCGAGTAATCTTTATCTCTGACGCTACAGCTAATCAAAAGAAATTTAGA 147
QY 408 AGACTTAATTAATTTGCTCTACGATTTGAACCTGAAAGTTTTCGATGGAGATTTAATCA 467
DB 148 TGAATTTCAAAAAGCTGCACCTAAAGCGAAATTTGTTATGTTGGTGCAGTGAAGAA 207
QY 468 AATATTTGATGCTTTTAAACCAATTTCAAAAGCTTTTAGGTAAAGAAAGAGGTTAAAA 527
DB 208 CTTAATTTGGTTCAATGAACAAACACCTGAAATATCGGAAATTTTACGA---TAAAG 264
QY 528 ACGCTTAGAAGAACACGATAAGAAATTTGAAGAAATATAAAAGAAATAACTATGGATA 587
DB 265 AGATAAGCTTAAGATTAATTAAGATTTAGATAACAAATTTGCTTCAATGAAGATA 324
QY 588 AATCAAAAGGATTTGCTCAGTAGCTGCTAATACAGGTTTCTGCTCATCCAGCAA 647
DB 325 AAGCAAAACCTTCAATAAAACCTGTTATGTTATTTACTAGTTTAAAGAGGTTGAATTA 384
QY 648 CTCTATGTTGGTCAATTTCTTAAGTCACTAGGTTTAAAGAGAGCTTAAGTCAATGTT 707
DB 385 ATTGGAACCTTAAAGGTCGTTTGGGATTTAGTTTACGATACATTTAGATTTCAATG 444
QY 708 TACTTAAAGGTTTAAAGTATCTTAAAGGACCTTACTTAAAGTAAAGCACTGAAACTTT 767
DB 445 TGATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 498
QY 768 ATCTCAAGTGAATCTGAGCGTATGTTCAATATG---ACAAAGAAAGAGTCTTAACGA 824
DB 499 TAATAAGAAATCCAGATGTTATTTTAGCGATGGATAGAGGTCAAGCGATAGTGGTAA 558
QY 825 ACCTTCACTAAAGAACTAGAAAAGATCTCTGTATGGAAGAAATTAACCGCTGTGAAA 884

Db 559 ATCAACTCGAAACCAAGCATTTAAATAATCTGTATTAAATAATGTTAAAGCAATTAAGA 618
QY 885 TCAACGTTGTTGATATTTTAGACCGTGAATTTATGGCAAGATCACGTGTTTAAATTTCTTC 944
Db 619 AGACAAAGTATATTAATTTAGATCTCTAAATTTATGTTACTTTTCAGCTGGATCACTACA 678
QY 945 AGAAGAAATGCAAAAGAACTTTGTTGAATTTCTTAAGAAAGATAGTAAAA 995
Db 679 TACAATTAACAAATTCGAGAACTTTGATAAAGTTGTAATAATTTTAAAA 729

RESULT 34
US-09-601-198-75/c
; Sequence 75, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 5652
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-75

Query Match 5.9%; Score 59.8; DB 4; Length 5652;
Best Local Similarity 44.4%; Pred. No. 0.0013;
Matches 241; Conservative 0; Mismatches 302; Indels 0; Gaps 0;

QY 225 AGATGTTAAACCTGTTGGGATAGCGATGATAACAAATAATCGTATTTTAAACCATTT 284
DB 3853 AAAAATTAATCGTATTCATACAGAAAGTTGTTCAAGATAAATATGTTAAATAATACATATA 3794
QY 285 AAGAGATAAATTTGAAAAATACACTTCTGTAGGACACGTAAGCAACCTTAACCTTAGAAGA 344
DB 3793 AGTTGTTAATCGTGAATACGAAGGTTATTTATGTAACCTGATCGAATTCCTTTATTTTCATT 3734
QY 345 AATCAGTAAACCTTAAACCCAGATTTTAAATTTATGCTGATATAATAGACACAAAGGTTATTTA 404
DB 3733 AGTTGCTCAATCTGATCTCTGCAATTTTAAACCAACAGGTTATTAATTTTAAATAATGTAAG 3674
QY 405 TAAAGACTTAAATAAATTTGCTCTTAAGTGAACCTGAAAGTTTTCGATGGAGATTATA 464
DB 3673 TACACACGAATATGACATCCAAACATTAACACATGAAAGATATTTAGTGATAGTAA 3614
QY 465 TGAAAAATTTGATGCTTTTAAACCAATTTCAAAAGCTTTAGTAAAGAGAGAGGTTAA 524
DB 3613 TGATTCAGTTATTTGTTGGAGGAAATTTGATTTCTAGATCGGAGTTAGTGATGAATCATATGT 3554
QY 525 AAAACGCTTAGAAGAACACGATAAGAAATTTGAAGAAATATAAAAGAAATAACTATGGA 584
DB 3553 AAATCGTAAAGCAATTTCAAGATTTTAAATGCACTGTTCTAGCGGAATTTACCTTTTAGAA 3494
QY 585 TAAAAATCAAAAGGTTATTTGCTGAGTAGCTGCTAAATCAGGTTTGTGCTCATCCAG 644
DB 3493 AACTGATGTTGATTAATATCCAAACAAAGAGTGGTCACTTTTAAATTTTTCATTAATAA 3434
QY 645 CAACCTTTATGTTGGTCAATTTCTTAAGTCACTAGGTTTAAAGAGAGCATTAAGTATGA 704
DB 3433 CGATCTTAAAAATCCAGTTTGAGAAACCAAAAGATATTTTGGTTCCGTTTAAATGCTGA 3374

QY 705 TGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAC 764
Db 3373 TGATCCCTAAAGCCCTTTTCTCTATAATAAAAAAAGCGTGGTTTTTTTACAAAAAATACGATGAATT 3314

QY 765 TTT 767
Db 3313 ATT 3311

RESULT 35
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESS: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match 5.9%; Score 59; DB 2; Length 19124;
Best Local Similarity 50.1%; Pred. No. 0.0024;
Matches 173; Conservative 0; Mismatches 170; Indels 2; Gaps 1;

QY 250 GATGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAAAAATACACT 309
Db 15570 GAATAAAAAATGAAAAAGATTATCAAAAAAATTTAAAAAATTTTATATATAAAAAA 15629

QY 310 TCTGTAGGACACGTAAGCACTTAATCTAGAGAAATCAGTAAACTTTAAACCATTTA 369
Db 15630 AATGATTATATAAAAAATATAAACAA--AAGAAAGAAAAAATTAATTAATAAAAAA 15687

QY 370 ATTATGCTGATAATAATAGACACAAAGGTATTTTATAAGACTTTAAATAAAATGCTCCT 429
Db 15688 AATATATATCAATAAACAACAAAAAATATATTTAAATAAATAAATATATATATATATC 15747

QY 430 ACGATTCACTGAAGAGCTTTCGATGGAGATTATTAATGAATATATTGATGCTTTTAAACA 489
Db 15748 ATAAAAATAAAAAAATTTAAAAAATGTTAAAAAATATATATACATAAAAAATAAAAA 15807

QY 490 ATTTCAAAAGCTTTAGCTAAAGAGAAAGAGGTAAAAAAGCGTTAGAGAACACGATAAG 549
Db 15808 ATTTATTAAATAAAAAAATTAATAATAAAAAATTTTAATTAATAAAAAAATAAT 15867

QY 550 AAAATTGAAGATATATAAAAAAGAAATTAATCTATGATGATAAAAATCAA 594
Db 15868 AATAAAAAAATTTAATTAATAAAAAAATTTAAAAAATTTAA 15912

RESULT 36
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JF2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 5.9%; Score 59; DB 4; Length 640681;
Best Local Similarity 49.6%; Pred. No. 0.005;
Matches 178; Conservative 0; Mismatches 180; Indels 1; Gaps 1;

QY 255 TAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAAAATGGAAAAATACACTTCTGT 314
Db 635938 TAAAAAAGAACTCTTTTATTAAAAAATATGATTTTAAATAGATATTTTCTTTAT 635997

QY 315 AGGAACACGTTAAGCACTTAAGAGAAATCAGTAAACTTAAACAGATTTTAATTAT 374
Db 635998 ATGAACCAATTAGATAAATTTAAAGAGAAATTAATTAATAAAAAAGAAATTGAACAATCTCG 636057

QY 375 TGCTGATAATATAGACACAAGGCTATTATAAGACTTAAATAAAATGCTCCTACGAT 434
Db 636058 AACAAATAATACAAAAAATTTATATTGCAAAAAAATAGAAATAGAGTTTGAAA 636117

QY 435 TGAACCTGAAAAGTTTCGATGGAGATTATAATGAAAAATTTGATGCTTTTAAAAACAATTTC 494
Db 636118 AAACAAGTATAAAAAAGTCAATTAATAATTTTCAGATTAATTTTTCGAAAAATAA 636177

QY 495 AAAAGCTTTAGTTAAAGAGAAAGGTTA-AAAAACGCTTTAGAGAAACACGATTAAGAAAA 553
Db 636178 AAGACGTTTGGAAAAAAGAAATACAATATATATTAAAGCAAAATAAATAGGATATTTTA 636237

QY 554 TTGAAGATATATAAAAAAGAAATTAATCTGGTATAAAATCAAAAGGTATTGCTGCAGTA 612
Db 636238 TATGATTTTTTTTTTAAAAAACAACGATAAAAAATTTATTAGGTATTGATTAGGTA 636296

RESULT 37
US-09-601-198-66
; Sequence 66, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Casseil, Gail H.
; APPLICANT: Chen, Ellison Y.

```
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 66
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-66

Query Match      5.8%; Score 58.8; DB 4; Length 861;
Best Local Similarity 44.8%; Pred. No. 0.0014;
Matches 225; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 251 ATGATAACAAAAAATCGTATTATTAACCCATTAAAGAGATAAAATGGAAAAATACATT 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 AGCTTAATGAAATATAACAATATTATGATAAAAAAGGTGATATATCCATAAAAAAGACAG 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 CTGTAGAACACGTAGACCACTTAACCTTAGAGAAATCAGTAACCTTAAACCCAGATTAA 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 266 ACATTACAAAAATTAAACCAATTAATCAACAAAAAGAAATTTAAAGGATTATTTAAAAA 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 371 TTATTCCTGATATAATAGACACAAAGGTATTATTAAGACTTAAATAAAATTTGCTCTTA 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 326 ATAAAGAGAAAAAATATAGGATCAATTAATAATTAATCTACGAAATCTTAATAACCAA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 431 CGATTGAACGTGAAAAGTTTCGATGGAGATTATAATGAAAATATGTGCTTTTAAAAACAA 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 386 TAATTAAATGTAATAAATGTAGATGATAAAAAACAAACGAAATTTCACTAAATTTAAAA 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 491 TTTCAAAAGCTTTAGTAAGAAGAGAGAGTTAAAGAGCTTAGAAGAACACGATAAGA 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 446 ATAATGATTATTATTTCTTAATGATAAAAAACAAATAAAATCAACGAAATTAATAATAA 505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 551 AAATTCAAGATATAAAGAAATAACTATGATGATAAAATCAAAAGGTATTGCTGCGAG 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 506 GTTATGAAGAAAAACCTTTAACTAAACGCTTCAATATTATTATGTAATTTGCGAAAA 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 611 TAGCTGTAATCAGGTTTGCTTGCTCATCCAGCAACTCTTATGTTGGTCAATTCCTAA 670
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 566 ATAAAGATTATACCAATTATTAGATTTCAAAACCTTATTTTCTAATGAAGAACATTAA 625
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 671 GTCACTAGGTTTAAAGAACATTAAAGTATGATGATGATGATGATGATGATGATGATGAT 730
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 626 ATGATATTAATTAATGAAAGCTTGTGTTTATTAATAATATCTATGAATAGTCAAAATCG 685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 731 TTAAGGACCTTACTTACAAAT 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 686 CTATAAGTAGTTTCAAGAAAT 707
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 38
US-09-949-016-14624/c
; Sequence 14624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14624
; LENGTH: 147382
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(147382)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14624

Query Match      5.6%; Score 56.8; DB 4; Length 147382;
Best Local Similarity 48.2%; Pred. No. 0.011;
Matches 193; Conservative 0; Mismatches 202; Indels 5; Gaps 1;

QY 196 TATTCATTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGAT 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137260 TATATAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 137201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 AACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATGGAAAAATACACTTCTGTA 315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137200 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 137141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 316 GGACACCTAGACCACTTAACCTTAGAAGAAATCACTAAACCTTAAACACGATTTAATTATT 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137140 TAAATACATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 137081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 GCTGATAATA-----ATAGACACAAAGGTATTTTATAAAGACTTAAATAAATTTGCTCTTA 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137080 TATAAATATATATATTAATATACATAAATAATGATATAAATAAATAAATAAATAA 137021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 431 CGATTGAACGTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAAACAA 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137020 TAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 136961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 491 TTTCAAAAGCTTTAGGTAAGAGAGAGGTAAAGAAAGCTTAAAGAAACGATAAGA 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136960 TATGTTAATATATACATATATAAATAAATAAATAAATAAATAAATAAATAAATAA 136901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 551 AAATTGAAGATATATAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136900 TAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 136861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 39
US-09-902-540-1318
; Sequence 1318, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1318
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(614)
; OTHER INFORMATION: unsure at all n locations
```

US-09-902-540-1318

```
Query Match          5.6%; Score 56.4; DB 4; Length 614;
Best Local Similarity 47.7%; Pred. No. 0.0042;
Matches 165; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

Qy 251 ATGATACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAAATACACTT 310
Db 180 ATAAAAAATAAAAAAATAAAAAAAGAAAAAGAAAAAACAACAACAAGAAAA 239

Qy 311 CTGTAGGAACACGTAAGCAACTTAACCTTAGAAGAAATCAGTAACTTAAACCAAGATTTAA 370
Db 240 AAGCGGATAAAAACAATAACACATATAACCAAAAAAGAACAAAAACAACA 299

Qy 371 TTATTGCTGATAATAATGACACAAAGGTATTATTAAGACTTAAATATAAATTTGCTCCTA 430
Db 300 AAAGAAAAAATAAGAAAGACAAAAAATAAAAAACAACAACAATAAAAAAATAAAAAA 359

Qy 431 CGATTGACTGAAAGTTTCGATGGAGATTATAATGAAAAATTGATGCTTTTAAACAA 490
Db 360 CAAAAACACACAAAAAACAATAAACAACAAAAAAGAAAAAATAAAAAAATAAAAAA 419

Qy 491 TTTCAAAAGCTTTAGGTAAGAAAGAGGTAAGAAAAACGCTTAGAAGAACACAGATAAGA 550
Db 420 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 479

Qy 551 AAATTGAAGATATATAAAAAAGAAATAACTATGGATATAAATAAATAAATAA 596
Db 480 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 525
```

RESULT 40

```
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28232)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (303398)..(303398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```

```
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (71352)..(71352)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855339)..(855339)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
```

Search completed: August 26, 2005, 07:02:59
Job time : 220 secs

```
Query Match 5.6%; Score 56.4; DB 4; Length 1664976;
Best Local Similarity 47.7%; Pred. No. 0.022;
Matches 204; Conservative 0; Mismatches 216; Indels 8; Caps 1;

Qy 314 TAGAACACGTAAGCAACCTTAAGAGAAATCAGTAACCTTAACCAAGATTTAATTA 373
Db 661285 TATTAAAGAAAGAAAGAAATTCGAAAGAAAGAAAGAAAGAAAGAGAGAGTTAGAGA 661226
Qy 374 TTGCTGATATATATAGACACAAAGGTATTTATTAAGACTTTAAATAAAATTTGCTCCTACGA 433
```

US-08-916-421B-1

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 00:31:21 ; Search time 4629 Seconds
(without alignments)
10551.492 Million cell updates/sec

Title: US-10-724-972a-2580
Perfect score: 1008
Sequence: 1 ggagtgggaatcagtgagagg.....gtaaaaagataaagtaa 1008

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1008	100.0	300698	1 AE016750	AE016750 Staphyloc
C 2	569	56.4	110000	1 BX571857_22	Continuation (23 o
C 3	569	56.4	304050	1 AP004829	AP004829 Staphyloc
C 4	569	56.4	307750	1 AP003136	AP003136 Staphyloc
C 5	569	56.4	347650	1 AP003364	AP003364 Staphyloc
C 6	567.4	56.3	110000	1 BX571856_23	Continuation (24 o
C 7	557.4	55.3	1044	6 AX622368	Sequence 10
C 8	336.2	33.4	2247	6 A64853	A64853 Sequence 10
C 9	269.2	26.7	298050	1 AP003132	AP003132 Staphyloc
C 10	269.2	26.7	347785	1 AP003361	AP003361 Staphyloc
C 11	263.8	26.2	975	6 AX621226	AX621226 Sequence
C 12	263.2	26.1	957	6 AX583924	AX583924 Sequence
C 13	262.8	26.1	110000	1 BX571857_10	Continuation (11 o
C 14	262.8	26.1	290150	1 AP004825	AP004825 Staphyloc
C 15	254.2	25.2	110000	1 BX571856_10	Continuation (11 o
C 16	222.6	22.1	668	6 AR354237	AR354237 Sequence
C 17	222.6	22.1	668	6 AR535793	AR535793 Sequence
C 18	190.8	18.9	5718	6 AX951891	AX951891 Sequence
C 19	190.8	18.9	37900	1 D86417	D86417 Bacillus su

C 20	190.8	18.9	197409	1 BSUB0005	Z99108 Bacillus su
C 21	168.4	16.7	110000	1 AE017333_11	Continuation (12 o
C 22	168.4	16.7	110000	1 CP000002_11	Continuation (12 o
C 23	166.6	16.5	12556	1 BSY14084	Y14084 Bacillus su
C 24	166.6	16.5	198743	1 BSUB0006	Z99109 Bacillus su
C 25	162.2	16.1	242	6 AR356438	AR356438 Sequence
C 26	162.2	16.1	242	6 AR537994	AR537994 Sequence
C 27	156.6	15.5	13792	1 AE006047	AE006047 Pasteurel
C 28	149.4	14.8	801	6 AX433651	AX433651 Sequence
C 29	146.4	14.5	301332	1 AE017012	AE017012 Bacillus
C 30	143.2	14.2	110000	1 AE017225_43	Continuation (44 o
C 31	143.2	14.2	110000	1 AE017334_43	Continuation (44 o
C 32	143.2	14.2	290029	1 AE017038	AE017038 Bacillus
C 33	140	13.9	110000	1 AE017355_43	Continuation (44 o
C 34	125.6	12.5	110000	1 BX571857_00	BX571857 Staphyloc
C 35	125.6	12.5	110000	1 BX571857_01	Continuation (2 of
C 36	125.6	12.5	290250	1 AP004822	AP004822 Staphyloc
C 37	125.4	12.4	2115	6 AR354486	AR354486 Sequence
C 38	125.4	12.4	2115	6 AR536042	AR536042 Sequence
C 39	124.6	12.4	110000	1 AE017355_06	Continuation (7 of
C 40	124.6	12.4	110000	1 CP000001_06	Continuation (7 of
C 41	124.2	12.3	304262	1 AE017005	AE017005 Bacillus
C 42	124	12.3	990	6 AX583941	AX583941 Sequence
C 43	124	12.3	999	6 BD237137	BD237137 Staphyloc
C 44	124	12.3	999	6 AR282355	AR282355 Sequence
C 45	124	12.3	3180	1 AF079518	AF079518 Staphyloc

ALIGNMENTS

RESULT 1
LOCUS AE016750/c
DEFINITION Staphylococcus epidermidis ATCC 12228, section 7 of 9 of the complete genome.
ACCESSION AE016750
VERSION AE016750.1
KEYWORDS GI:27316220
SOURCE Staphylococcus epidermidis ATCC 12228
ORGANISM Staphylococcus epidermidis ATCC 12228
REFERENCE 1 (bases 1 to 300698)
AUTHORS Zhang,Y., Ren,S., Li,H., Fu,G., Lu,L., Lu,G., Jia,J., Tu,Y., Qin,Z., Chen,Z. and Wen,Y.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES
source

300698 bp DNA linear BCT 01-JAN-2003
Staphylococcus epidermidis ATCC 12228, section 7 of 9 of the complete genome.
Staphylococcus epidermidis ATCC 12228
Bacteria; Firmicutes; Bacillales; Staphylococcus.
1 (bases 1 to 300698)
Zhang,Y., Ren,S., Li,H., Fu,G., Lu,L., Lu,G., Jia,J., Tu,Y., Qin,Z., Chen,Z. and Wen,Y.

1..300698
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="taxon:176280"
complement(114..186)
/product="tRNA-Lys"
complement(190..264)
/product="tRNA-Gln"
complement(269..352)
/product="tRNA-Tyr"
complement(369..444)
/product="tRNA-Val"
complement(451..522)
/product="tRNA-Glu"
complement(524..598)
/product="tRNA-Asn"
complement(612..726)
/product="5S ribosomal RNA"
complement(886..3713)
/product="23S ribosomal RNA"
complement(2718..2831)
/locus_tag="SEI755"


```

CDS      complement (2718..2831)
          /locus_tag="SE1755"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="AAO05396.1"
          /db_xref="GI:27316221"
          /translation="MMIIGGRALFGAPLGPLPNSDKLRMPINLTWSEHG"
          /complement (3977..5530)
          /product="16S ribosomal RNA"
          /complement (5985..7052)
          /locus_tag="SE1756"
          /complement (5985..7052)
          /locus_tag="SE1756"
          /codon_start=1
          /transl_table=11
          /product="ATP-binding Mip-like protein"
          /protein_id="AAO05397.1"
          /db_xref="GI:27316222"
          /translation="MLTVNQKEIVGALKDPIIDVPLRESEGIVDVSIKONINHVSK
          VAMAQGGQQLDQLMAIVKALKENGANTVGIREFEELPSEVVERIYIGKSKPTIE
          LLSQNPVEFISIASGGVGKSTAVNLAVALAREGKGVLDADYGFSPDMMGI
          DERQIDGKEIIPVERHGKVIISMAFVEENAPVIRGPMGLGKMLTNPFTEVQWGEID
          YLLDLPGCTGDALDVHSLMPPSKEIIVTTPHTAFAVAARAGAMAKHTHTILGVI
          ENMYPSEKTEGKEVYFGKGGKLSDELETQLFAELPLEQPTWPNDFSPSIQSD
          DRLELYLLIARKVIATQKQ"
          /complement (7203..8642)
          /locus_tag="SE1757"
          /complement (7203..8642)
          /locus_tag="SE1757"
          /codon_start=1
          /transl_table=11
          /product="multidrug resistance protein"
          /protein_id="AAO05398.1"
          /db_xref="GI:27316223"
          /translation="MASVELSKTRNLIVSMLSAFVAILNOTLLNTALPHIMRELN
          INENSWLITGFMVGVMTPLPAYLMDKVTPLVYLGAMGSLGISVAIAIPNEG
          VIMTARVQAGAILPMLQPTLTFLFSKHRGFAMGLAGLVIQPAPIGPTTGLI
          IDNASRVPFIVIGIALVTFIFGVFSSISNTTKETKLDKRSVISTYLGFLMLYAF
          SSAQNLGSPNLSILSIIGIFVKRQITISNPLLNKIKNKIFCFSTISMI
          IMLMVGPAIIPLYVQNALGSLGLVMPGAIINGIMSVFTGKFDYGRPLI
          FTGILLISCTFLCKVDTSYTVLIVYAIRMPAVSLMMPINTTGINALKTEDIS
          HGTAIMRGRVMAQSLCTALAVTMSIGAQVWSSESSEHASKEMIQRSVAVGVDSFA
          LVTVFVIAFVLAIFKEPKNHLTNSRKV"
          /complement (8912..9382)
          /locus_tag="SE1758"
          /complement (8912..9382)
          /locus_tag="SE1758"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO05399.1"
          /db_xref="GI:27316224"
          /translation="WMKFLNKSVHLVTLVLTIVTIVISGAIFLTFLGFLYGLSLRIL
          IYHLGDSYNGKFDNLIYVGSIVLGYFTLFSIEHLMDFKNLPKNPFGQINFH
          LISIVTITMFIYVHIHYVNHFWIMIIIGFLVCKEVFTPESKNLANKK"
          /complement (9570..10907)
          /locus_tag="SE1759"
          /complement (9570..10907)
          /locus_tag="SE1759"
          /codon_start=1
          /transl_table=11
          /product="multidrug transporter"
          /protein_id="AAO05400.1"
          /db_xref="GI:27316225"
          /translation="MNLKSIITVMALILMFMAAETSIISIALPTIKNSLNAGNLVS
          LVTVYFALVIANPIVIGELMSRKLIIVAVGVLLPALGSLMGLSQTPTFLISRT
          WQFGAGVMSLSQIVPKLAFEPIDRYKIMGVGSVWGSIISIFGLIGALLERFASWH
          VLFYNIDPVIATVILMTTFHPDETQVQSRREDIKGLIIFIFIALMGLLNQHH
          IIPNIFSLALVMLLPKFIENSIEQPLTKFNISIVLVFITDLIAITLMGYNL
          YIPVYQKLSLPSQSGFVIFPLSAVMTIFNFAKTEAHTKRLTYICFFVLLVS
          SLMTMFGKLPLLIATAFVAVFAGLFGYIYTKDSIVIQEETSPKNNKMMSFALTQNL
          GSSVGSITMGVYALNVGLFGSNLHNVLGLVLIITAVCLIVMMMTLYKSNITIQS"

gene      complement (11146..11829)
          /locus_tag="SE1760"
          /complement (11146..11829)
          /locus_tag="SE1760"
          /codon_start=1
          /transl_table=11
          /product="hemolysin III"
          /protein_id="AAO05401.1"
          /db_xref="GI:27316226"
          /translation="MSQSKRKNDSSVVETFKDIIPLTFGEIGNAASHGAALITLFI
          LPYAAVHSFNNGGTLESISVSIVYISFMFISSTIYHSMQNTTSKYILRIIDHSMI
          YVAISGTYTPVLLSVGVGMLGWLVTILLGTLTWGLIKYSIATKVNRLSLIVLVLMG
          WVGIIPLPIIIMRTSWMFIFPFLGGLSYIGAMFYAQKNKPYFHHIWHIIVLASPL
          HMGIGFYFM"
          /complement (11848..13035)
          /locus_tag="SE1761"
          /complement (11848..13035)
          /locus_tag="SE1761"
          /codon_start=1
          /transl_table=11
          /product="UDP-N-acetylglucosamine pyrophosphorylase"
          /protein_id="AAO05402.1"
          /db_xref="GI:27316227"
          /translation="MLDKNOLKYNQOHLVEYKELMSNEKNALDEKVDQNLAEIQD
          LYODLVNRTIDDVSVSVKVKSRLANEERHTYEQGYEARNGERPAVLMMAGG
          QGTRLGKYGKGSFEISGTSLSLFEQAQQLRLKEETGHTINWIYIMTSDINHKTIEYF
          KQHYFNYDANHIFPKQDNIVALSESGKVLNKRGDHIMETPNGNGGVFKSKKAGYL
          DKMOQHVKYIFLNDINLVKLDPLFAGTFTVQSKDITSKTLPQKDSGVRLNVV
          DCKDTVLEYSLELTDIVQNFNANIGIHAPKLGFIISAVDRELPHYLAIKQLQDEN
          FGVRPTLKFELFYDIPRYGTSFTVLQVPREBEFPLKNKEGKDSVHTATEDLKRK
          DLI"
          /complement (13090..13599)
          /locus_tag="SE1762"
          /complement (13090..13599)
          /locus_tag="SE1762"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO05403.1"
          /db_xref="GI:27316228"
          /translation="MTGKTHASGCLVGAIITQVPHFTDIFTISIVLSVSIILPDI
          CHTQSKIGRFRILTSFFVRILFGHRTFTSHLLFIIGISFLYFTQTPMYNVAIVIGM
          FSHVLDILPRGVKLLYPLPFNIVSPHFKTGLVDVSLATALSVGAIYTLFPQYLN
          TMHYWLK"
          /locus_tag="SE1763"
          /locus_tag="SE1763"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO05404.1"
          /db_xref="GI:27316229"
          /translation="MAMTVKNDNEVRIOQRVADIKIPNNKIKNTQDDIHAVEPEN
          GKEISRIGSYFGKTNRVLIDTDQHLIYIITQNDQKVYNELTK"
          /complement (14189..14503)
          /locus_tag="SE1764"
          /complement (14189..14503)
          /locus_tag="SE1764"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="AAO05405.1"
          /db_xref="GI:27316230"
          /translation="MFVVTNRTVTKGYAKQMAPNFTKGGPIESLKGFEIEVMQIDK
          DVSSEDMYVNSWETEDOPKNNVNSDVFKQAHKNTGKSEDSPIKSVIKSNVLSL
          RR"
          /complement (14682..16055)
          /locus_tag="SE1765"
          /complement (14682..16055)
          /locus_tag="SE1765"
          /codon_start=1
          /transl_table=11

```

/product="conserved hypothetical protein"	
/protein_id="AAO05406.1"	
/db_xref="GI:27316231"	
/translation="MNGPKYDIIVRGRIKFTMSLLGLILFLVPIPVVDGKQQTLLP	
LAPLAGLWGLMGLIPILIVITIVSGILITLCSITITKNGKLNFGGLMSSAFNFKIGW	
LVRLVLAFLKDFLPLNIPGMEIKSEDTGGLVFSLLPLVAVFLFAAFLPLMLMEYG	
LLELGPFRIMRPLFTLPGRSVDNLASFIGDGTGVLITSQYGEYYSREAVT	
ISTFVSIVITFAIVIAETIRMQDOFFYFLITVVISLIIAAMIPRIWPLKNIPEVA	
KEVSEARNELPEKTKALYGFPLATEVGIKSPGKFLSLGPKTVVDMVFLPLPVV	
MSIGETIATLIANYTPVEIIGKPFVPLELLQIPEAHESOTILIGPADMFLPSLIE	
GVQNDVTRFVIGALISIVLSEVGGVILGSKIPVSIKLFMFLPLRIITLPIIAL	
LAHLFIG"	
complement (16352..17317)	
/locus_tag="SE1766"	
gene	
Query Match 100.0%; Score 1008; DB 1; Length 300698;	
Best Local Similarity 100.0%; Pred. No. 1.8e-144; Indels 0; Gaps 0;	
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CGAGTGGGAATCAGTCAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTGTTTAA 60
Db	19364 GGAGTGGGAATCAGTCAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTGTTTAA 19305
Qy	61 ATTGCAACTGCAAGCATGTGGAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
Db	19304 ATTGCAACTGCAAGCATGTGGAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 19245
Qy	121 GATGAGTTGAATCAAGCACCAAGAGGTACTAGCAAGTACCTTAAACCCCTAAACGT 180
Db	19244 GATGAGTTGAATCAAGCACCAAGAGGTACTAGCAAGTACCTTAAACCCCTAAACGT 19185
Qy	181 GTTGTGTTCTTGGATGATTCAATTTGTTGATGGTTAGTTGCTTTAGATGTTTAAACCTGTT 240
Db	19184 GTTGTGTTCTTGGATGATTCAATTTGTTGATGGTTAGTTGCTTTAGATGTTTAAACCTGTT 19125
Qy	241 GGGATAGCGGATGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGA 300
Db	19124 GGGATAGCGGATGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGA 19065
Qy	301 AAATACACTTCTGTAGGAAACAGTAAAGCAACCTTAAGTAAAGAAATCAAGTAAACTTAA 360
Db	19064 AAATACACTTCTGTAGGAAACAGTAAAGCAACCTTAAGTAAAGAAATCAAGTAAACTTAA 19005
Qy	361 CCAGATTTAAATTTGCTGATTAATATACACAAAGGTATTATTAAAGACTTTAAATAAA 420
Db	19004 CCAGATTTAAATTTGCTGATTAATATACACAAAGGTATTATTAAAGACTTTAAATAAA 18945
Qy	421 ATTGCTCCTACGATTGAACCTGAAAGTTTCGATGGAGATTATTAATGAAAAATTTGATGCT 480
Db	18944 ATTGCTCCTACGATTGAACCTGAAAGTTTCGATGGAGATTATTAATGAAAAATTTGATGCT 18885
Qy	481 TTTAAAAACAATTTCAAAAGCTTTAGTAAAGAAAGAAAGGTAAAGAAAGCTTAGAGAA 540
Db	18884 TTTAAAAACAATTTCAAAAGCTTTAGTAAAGAAAGAAAGGTAAAGAAAGCTTAGAGAA 18825
Qy	541 CACGATAAGAAATTCAGCAATATAAAGAAATAACTATGATATAAATAATCAAAAGGTA 600
Db	18824 CACGATAAGAAATTCAGCAATATAAAGAAATAACTATGATATAAATAATCAAAAGGTA 18765
Qy	601 TTGCTGCGAGTAGTCTCTAAATCAGGTTTTCCTTGTCTCATCCAGCAACTCTTATGTTGT 660
Db	18764 TTGCTGCGAGTAGTCTCTAAATCAGGTTTTCCTTGTCTCATCCAGCAACTCTTATGTTGT 18705
Qy	661 CAATTCCTAAGTCAACTAGTCTTTTAAAGAGCAATTAAGTGTATGATGATGATGATGATGAT 720
Db	18704 CAATTCCTAAGTCAACTAGTCTTTTAAAGAGCAATTAAGTGTATGATGATGATGATGATGAT 18645
Qy	721 AGTAAGTATCTTAAAGCACTTACTTACAAATGCAACACTGAACTTTTATCTCAAGTGAAT 780
Db	18644 AGTAAGTATCTTAAAGCACTTACTTACAAATGCAACACTGAACTTTTATCTCAAGTGAAT 18585
Qy	781 CCTGAGCGTATGTTTCAATATGACAAACAAAGCAAGTTTCTAAACCACTTCACTTAAAGAA 840

Db	18584	CCTGAGCGTATGTTTATAATGACAAACAAAGCAAGTTCTTAACGAACCTTCACTAAAGAA	18525
Qy	841	CTAGAAAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTATATT	900
Db	18524	CTAGAAAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTATATT	18465
Qy	901	TTAGACCGTGAATTTATGGGCAAGATCAGCTGGTTTAAATTTCTTCAAGAAATGGCAAAA	960
Db	18464	TTAGACCGTGAATTTATGGGCAAGATCAGCTGGTTTAAATTTCTTCAAGAAATGGCAAAA	18405
Qy	961	GAACTTGTGAATTTATCTAAGAAAAAGATAGTAAAAAAGATAATAAGTAA	1008
Db	18404	GAACTTGTGAATTTATCTAAGAAAAAGATAGTAAAAAAGATAATAAGTAA	19357
RESULT 2			
BX571857.22/c			
WPCOMMENT			
Sequence split into 28 fragments LOCUS BX571857 Accession BX571857			
		Fragment Name	Begin End
		BX571857_00	1 110000
		BX571857_01	100001 210000
		BX571857_02	200001 310000
		BX571857_03	300001 410000
		BX571857_04	400001 510000
		BX571857_05	500001 610000
		BX571857_06	600001 710000
		BX571857_07	700001 810000
		BX571857_08	800001 910000
		BX571857_09	900001 1010000
		BX571857_10	1000001 1110000
		BX571857_11	1100001 1210000
		BX571857_12	1200001 1310000
		BX571857_13	1300001 1410000
		BX571857_14	1400001 1510000
		BX571857_15	1500001 1610000
		BX571857_16	1600001 1710000
		BX571857_17	1700001 1810000
		BX571857_18	1800001 1910000
		BX571857_19	1900001 2010000
		BX571857_20	2000001 2110000
		BX571857_21	2100001 2210000
		BX571857_22	2200001 2310000
		BX571857_23	2300001 2410000
		BX571857_24	2400001 2510000
		BX571857_25	2500001 2610000
		BX571857_26	2600001 2710000
		BX571857_27	2700001 2799802
Continuation (23 of 28) of BX571857 from base 2200001 (BX571857 Staphylococcus aureus str			
		Query Match	56.4%; Score 569; DB 1; Length 110000;
		Best Local Similarity	73.3%; Pred. No. 1.1e-77;
		Matches	728; Conservative 0; Mismatches 265; Indels 0; Gaps 0;
Qy	7	GAATCAGTGAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTTGTATTTGCA 66	
Db	48085	GGACGATGAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTTGTATTTGCA 48025	
Qy	67	ACTGCAGATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGCA 126	
Db	48025	GTTGCAGCTTGTGGTAAATTCAGTAAAGAAATCAATCAACTAAGATATCT 47966	
Qy	127	GTGGAATCAAGCACGAAGAGGTACTACGAAAGTACCTTAAACACCCCTAAACGTTGTTT 186	
Db	47965	ATTTCCGTAAGAGTGAATGGTACAGTAAAGTACCTTAAAGATGCAAAAGATCGTT 47906	
Qy	187	GTTCTTGAGTATTCATTTTGTGATCGGTTAGTTGCTTTAGATGTTTAAACCTGTGGATA 246	
Db	47905	GTTATTAGTACTCATTTTGCAGATGCAATTAGCAGCATTAAGCGTTAATCCAGTTGTTAT 47846	
Qy	247	GGGATGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAATATAC 306	
Db	47845	GCTGATGATGGTGAAGAAAAACGTATCAATTAACCCAGTTAGAGAAAAATTTGGGATTAT 47786	

QY	307	ACTTCTGTAGCAACGTAAGCAACCTAACTTAGAGAAATCAGTAACACTTAACCCAGAT	366
Db	47785	ACTTCTGTAGTACACGTAAGCAACCTTAGAGAAATTAGTAATTAACCCGAT	47726
QY	367	TAAATTAATTCGTGATAAATAAGACACAAAGTATTATTAAGACCTTAATAATAAATTGCT	426
Db	47725	TAAATTAATTCGTGATAGCAGTAGACATTAAGGTATTATTAAGAAATTAAACAAATTGCA	47666
QY	427	CCTACGATTGAATGAAAAGTTTCGATCGAGATTATTAATGAATAATTAATGATGCTTTTAAA	486
Db	47665	CCAACATTATCATTAAGAGTTTTCGATCGAGACTACAAACAAACATTAAATTCGTTCAA	47606
QY	487	ACAATTTCAAAGCTTTAGTTAAAGAAAGTAAAGGTAAGAAACGCTTTAGAGAACACAT	546
Db	47605	ACAATTCGTAAGCTTTAAATAAGAAAGAAAGGCGAAACGCTTCCTGAGCATGAT	47546
QY	547	AAGAAAATTGAAGATATATAAAGAAATAACTATGGATAAAATCAAAAGGTATTCGCT	606
Db	47545	AAATTAATCAATAAGTATAAGATGAATTAATTTGATAGAAATCAAAAGGTGCTTCA	47486
QY	607	GCAGTAGCTGCTAAATCAGGTTTGCTCATCCAAAGCAACTCTTATGTTGGTCAATTC	666
Db	47485	GCAGTAGTTGCTTAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGCAATTT	47426
QY	667	CTAAGTCAACTAGGTTTAAAGAACATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG	726
Db	47425	TAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGAACAAAAGGTTTAAAGTAA	47366
QY	727	TATCTTAAAGGACCTTACTTCAAAATGAACACTGAACTTTATCTCAAGTGAATCCTGAG	786
Db	47365	TATTTGAAGGACCTTACTTCAATAGACACTGAACTGAACTTGAATTAATCCAGAG	47306
QY	787	CGTATGTTTATATGACAAACAAAGCTTCTTAACGAACCTTCTACTAAAGAACTAGAA	846
Db	47305	CGTATGATCATTTATGACTGATCACGCTAAAGAAATTCGCTGAATTAAGAAATTAACA	47246
QY	847	AAAGATCCTGTATGAAGAAATTAACCGTGTGAAGAAATCAACGTTGATATTTAGAC	906
Db	47245	GAAGATGCACATCGAAGAAATTTGAATGCAGTTAAAGAAATTAATCGCGTGATATTGTC	47186
QY	907	CGTGACTTATGGCAAGATCACGCTGGTTAAATTTCTTCAGAGAAATGCGCAAAAGACTT	966
Db	47185	CGTGATGTTTGGGCAAGATCTCGTGCTTAATTTCTTCGAAGAAATGCGTAAAGACTT	47126
QY	967	GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT	999
Db	47125	GTTGAATTATCAAAAAAGAAACAAAGTAAGT	47093
RESULT 3			
AP004829/c			
LOCUS			
Staphylococcus aureus subsp. aureus MW2 DNA, linear BCT 20-DEC-2002			
strain:MW2, section 8/10.			
AP004829.BA000033			
AP004829.1 GI:21205117			
Staphylococcus aureus subsp. aureus MW2			
Staphylococcus aureus subsp. aureus MW2			
Bacteria; Firmicutes; Bacillales; Staphylococcus.			
1			
Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A.,			
Nagai, Y., Iwama, N., Asano, K., Naimi, T., Kuroda, H., Cui, L.,			
Yamamoto, K. and Hiramatsu, K.			
Genome and virulence determinants of high virulence			
community-acquired MRSA			
Lancet 359 (9320), 1819-1827 (2002)			
22040717			
12044378			
2 (bases 1 to 304050)			
Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,			
Nagai, Y., Asano, K., Iwama, N., Baba, T., Kuroda, M., Hiramatsu, K. and			
Kikuchi, H.			
Direct Submission			
Submitted (06-MAR-2002) Director-General, Biotechnology Center,			
National Institute of Technology and Evaluation, Biotechnology			
Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan			
(E-mail:bio@nitech.go.jp, URL:http://www.bio.nitech.go.jp/,			
Tel:81-3-3481-1933, Fax:81-3-3481-8424)			
Location/Qualifiers			
1. .304050			
/organism="Staphylococcus aureus subsp. aureus MW2"			
/mol_type="genomic DNA"			
/strain="MW2"			
/db_xref="taxon:196620"			
complement(629..1072)			
/gene="MW1947"			
complement(629..1072)			
/gene="MW1947"			
/note="ORFID:MW1947			
hypothetical protein, similar to acetyltransferase			
[Genomic island nu Sa 4mw]"			
/codon_start=1			
/transl_table=11			
/protein_id="BAB95812.1"			
/db_xref="GI:21205118"			
/translation="MSIITRLFNNSDFEKLNLCKLYDDLGYPNTNNDLKKRLKKITN			
HDDYLLILLIKENKIIGLSGCKMFMFEKNAEYMRILAFVIHSEFRKKGKGRLLADS			
EFSKRLNCKAITLNSGNRNLRLSAHLKLYSDNGVYVSTSGTKQL"			
complement(1178..1696)			
/gene="MW1948"			
complement(1178..1696)			
/gene="MW1948"			
/note="ORFID:MW1948			
[Genomic island nu Sa 4mw]"			
/codon_start=1			
/transl_table=11			
/product="hypothetical peroteine"			
/protein_id="BAB95813.1"			
/db_xref="GI:21205119"			
/translation="MDNINKKEKRVLSITAITISFILNFGFLTSLAWSINSMGKQ			
CISNLSDDTFAVISVMTLTGTAIGGIAPSIYGIPTSIDKSSFRASIDWQEQ			
RNNYKDLMLLPFTLLSLIVIFIQFAYSFKNTAIIIFIILVVMFFILLICYK			
ITVMICRYFSKE"			
complement(1932..2522)			
/gene="MW1949"			
complement(1932..2522)			
/gene="MW1949"			
/note="ORFID:MW1949			
hypothetical protein, similar to bacteriophage terminase			
small subunit [Genomic island nu Sa 4mw]"			
/codon_start=1			
/transl_table=11			
/protein_id="BAB95814.1"			
/db_xref="GI:21205120"			
/translation="MSELTAQARFVNEYIRTLNVTQSAIKAGYNANSAHTACRLLK			
KPHIKOYIOBKDKIIDENLVLTAKELLHLVLTATVDETETKEVVVKGEYKVPQSG			
KVQLYNHEVELIIVPIKPSDLKAREMLGKYLFTDKHDNGNVPIFINIGWDGG			
DEGLDKAVDVSNDNPNTIVDDIPLEDYESIRFL"			
complement(2519..2734)			
/gene="MW1950"			
complement(2519..2734)			
/gene="MW1950"			
/note="ORFID:MW1950			
[Genomic island nu Sa 4mw]"			
/codon_start=1			
/transl_table=11			
/product="truncated hypothetical protein"			
/protein_id="BAB95815.1"			
/db_xref="GI:21205121"			
/translation="MWTDDYNDHIKKNHIEIVDLATDKFEYIFDSLIQGLSYODI			
SSALDCSEQSVILWYETILDKIVGVIK"			
2864..3338			
/gene="truncated int"			
2864..3211			

```

/gene="truncated int"
/Note="ORFID:MW1951
truncated hypothetical protein, similar to integrase
[Genomic island nu Sa 4mw]"
/codon_start=1
/transl_table=11
/protein_id="BAB95816.1"
/db_xref="GI:21205122"
/translation="MKWSDIDFENNTISITKYNNPNKKKFOILPPKTESSIGKIS
VDPNPKVLRDKYNQNNKKNELYNHCHNEKINVDNTNHYVYQHTQKLYALIQI
HTLRVLRIRCTH"
3072..3338
/gene="truncated int"
/Note="ORFID:MW1952
truncated hypothetical protein, similar to integrase
[Genomic island nu Sa 4mw]"
/codon_start=1
/transl_table=11
/protein_id="BAB95817.1"
/db_xref="GI:21205123"
/translation="MITIVMKKLSMWIOTIMSRISILKNISHTSFYTHCALLIESGV
HIKEQERLRHKDINTWNIYAKITNSYKXDAHKFSHQMEDVS"
complement(3433..5049)
/gene="groEL"
/complement(3433..5049)
/gene="groEL"
/Note="ORFID:MW1953"
/codon_start=1
/transl_table=11
/product="GroEL protein"
/protein_id="BAB95818.1"
/db_xref="GI:21205124"
/translation="MVKOLKSEDAQALRGVQOLANAVKVTIGPKGRNVLDKEFT
APLINDGTIAKELEDPYENWGAQLVQEVANKTNEIAGDGTATVLAQMIQEG
LKNVTSGANVGLRGIDKAVKVAVEALHENSQKVENKNEIAQVGAISAADEEIGRIY
SEAMKGVNDGVITTEESNGLNEVEVQFDRGYSPYVMTDSKXVAELERPVI
LVTDKISFOIDILPELVQVQSNRPILIVADEVEDGALTNILNMRGTFTAVAKA
PGFGRRRAMELDAALGAOVIITDGLDLKDAISIDMLGTASKVEVTKDNTVVDDG
GDENSIDARVSLQSKIETESDSDPREKLOELAKLGGVAIKVGAASETELKERKL
RIEDALNSTRAAVEGI VAGGGTALVNTQKVEISEABEGDIETGYNIVL KALTAPVQ
AMFLTTEAVVASIPEKNDQNMGGPMGM"
complement(5125..5409)
/gene="groES"
/complement(5125..5409)
/gene="groES"
/Note="ORFID:MW1954"
/codon_start=1
/transl_table=11
/product="GroES protein"
/protein_id="BAB95819.1"
/db_xref="GI:21205125"
/translation="MLKPTGNRVIIIEKKEQETTKSGIVLTDSAKEKNEGIVAVGT
GRLLNDGTRVTEVKEGDRVVFQVAGTEVKREDNETYLVLNEEDILAVIE"
5584..6327
/gene="MW1955"
5584..6327
/gene="MW1955"
/Note="ORFID:MW1955"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95820.1"
/db_xref="GI:21205126"
/translation="MTRLWASLLTVIYILSQFLPLIIVKLLPFVQYSGIETKAVIY
IOLVFLIAATIIILINKIKNPTELEVEPKKYIIPWALLIGFALVMYQWVSIY
LTQVYGGQVSPNTEKIIIAIKPIPIFFIISIGLLEEVYFRKVIIFGELFNATKGN
RVAFIATVTSLEIFALAHNDPKFIPVYFGMGIIFSLAVVWTKRLAVPIIHHMQNG
FVIFOLLNPEALUKATEQANFIYHIFIP"
complement(6352..7611)
/gene="MW1956"
complement(6352..7611)
/gene="MW1956"

```

```

/Note="ORFID:MW1956
hypothetical protein, similar to SdrH"
/codon_start=1
/transl_table=11
/protein_id="BAB95821.1"
/db_xref="GI:21205127"
/translation="MSYHFWFKMLLSTLILSSSSGLATHTVKADNLNGEKPTTN
LNHNITSPVSNSENNNETGTTPHESNOTEGTGSNRDANPDSNNVKKPDSNNQNPST
DKPDPNPNQSPNPKPDPKPDPPKPNPDPKPNPDPKPNPDPKPNPDPKPNPDP
DKPDPNPKPDPKPNPDPKPNPDPKPNPDPKPNPDPKPNPDPKPNPDPKPNPDP
ASDGSOGQPNQNGNSONTGNTFVSORFALANGAYKYNPVLNKLKGLADYIG
EVTDEDIYIIRKQNFSGNALGLOQSNVFRFOYFNPLKSEYRYRLNDRVOLALIT
GEIGSPDLKPEDKPDSPKQSPPEPHEKDDFTVVKKQEDNKKASATSKSLAIVCS
MMVVFSLFLFVKNKKKNKQESRR"
7808..8434
/gene="MW1957"
7808..8434
/gene="MW1957"
/Note="ORFID:MW1957"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95822.1"
/db_xref="GI:21205128"
/translation="MGLFKDKKAMTFNMEERRSYNLKDSISISDDELESMTAHA
VKHVPSSQSSTRIVLLNNDNNKFWDNTHKILKEVMGENDREFTQKIDNFKHSY
GTILFYEDQDVYGLQOMPNYIDNFATWSTQTNMHQFPAITWALTATKIGIGASLOHYN
PLVDEMTSNEFNIPKSWKLIQAQMPFGDIREAAGEKTFNVEDRFVIK"
8795..9580
/gene="MW1958"
8795..9580
/gene="MW1958"
/Note="ORFID:MW1958"
Query Match 56.4%; Score 569; DB 1; Length 304050;
Best Local Similarity 73.3%; Pred. No. 8.6e-78; Indels 0; Gaps 0;
Matches 728; Conservative 0; Mismatches 265;
QY 7 GAATCAGTGAGAGGTTTAAAAATTTAAAGTGTAAATTTGGCTTATTTGTTTGTGTTTAAATGCA 66
DB 171480 GGAAGATGAGAGGTCTAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 171421
QY 67 ACTCAGCATGTGGAAATAAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGA 126
DB 171420 GTTGCAGCTTGTGGTAATACGGATAATTCACGTAATAAAGAAATCATCACTAAAGATACT 171361
QY 127 GTTGAATCAACGACGAGAGGTACTACGAAAGTACCTTAAACACCCCTAAACCGTGTGTT 186
DB 171360 ATTTCCGTAAGAAGATGANAATGGTACAGTAAAGATACCTTAAAGATGCAAAACGATCGT 171301
QY 187 GTTCTTGAGTATTTCATTTGTTGATCGTTAGTTGCTTTAGATGTTTAAACCTGTGGGATA 246
DB 171300 GTATTAGAGTACTCAITTTGCAGTGCATTAGCAGATTAGAGGTTAATCCAGTTGTTAT 171241
QY 247 GCGGATGATAACAAAAAATTCGTATTATTAACCATTAAGAGATAAAATTTGGAATATAC 306
DB 171240 GCTGATGATGTAAGAAAAAAGCTATCATTAACCAAGTTAGAGAAAAAATTTGGGATTAT 171181
QY 307 ACTTCTGTAGGAACAGGTAGCAACCTTAATTAGAGAAATCAGTAACTTAAACCGAT 366
DB 171180 ACTTCTGTAGGTACACGTAAACAGCCAACTTAGAGAAATTTAGTAAATTTAAACCGGAT 171121
QY 367 TTAATTTATGCTGATAATAATAGACACAAAGGTATTATTATAAGACCTTAATAATTTGCT 426
DB 171120 TTAATTTATGCTGATGACGATGAGACATTAAGAGGTATTATTATAAGAAATTTAAACAAATTTGCA 171061
QY 427 CCTACGATTGAATGAAAAAGTTTCGATGGAGATTATAATGAAAAATTTGATGCTTTTAAA 486
DB 171060 CCAACATTATCATTAAGAGTTTTTGTATGAGACTACAAACAAAAACATTAATTTGTTCAA 171001
QY 487 ACAATTTCAAAGCTTTAGGTAAAGAGAGAGGTAAAGAAACCGTTAGAGNACACGAT 546
DB 171000 ACAATTTGCTTAAAGCTTTTAAATAAGAAAAAGAGGCGGAAAAACGCTCTTCTGAGCATGAT 170941

```

Qy	547	AAGAAATTGAAGATATATAAAGAAATAAATACTATGGATAAAATCAAAAGTATTGGCT	606	gene	/sub species="aureus N315"
Ds	170940	AAATTAATCAATAGTATAAGATGAAATTAATTTGATAGAAATCAAAAAGTGTCTCCA	170881	CDS	/db_xref="taxon:158879"
Qy	607	GCAGTAGCTGCTAAATCAGGTTTGCTCTCATCAAGCAACTCTTATGTTGGTCAATTC	666		297..1907
Ds	170880	GCAGTAGTTGCTAAAGCTGGTTTATTAGCATATCAAACTATTCATATGTTGGACATTT	170821		/gene="SA1853"
Qy	667	CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGTATGTTACTTAAAGGTTTTAAGTAAG	726		/note="ORFID:SA1853
Ds	170820	TTAAACGAACTAGGATTTAAAAATGCAATTAAGTGACGATGAACAAAAGGTTTAAGTAAA	170761		hypothetical protein, similar to DNA mismatch repair
Qy	727	TATCTTAAAGACCTTACTTCAAAATGAACACTGAACCTTTATCTCAAGTGAATCCTGAG	786		protein MutS"
Ds	170760	TATTTGAAAGACCTTACTTCAATTAAGCATGCACTGAACCTTTAGCTGATTTAAATCCAG	170701		/codon_start=1
Qy	787	CGTATGTTCAATAAGACAAACAAAGCAAGTCTTCAAGCAACCTTCACTAAAGAACTAGAA	846		/transl_table=1
Ds	170700	CGTATGATCATATGATGATCACGCTAAAAAGATTTCTGCTGAATTCAGAAATTACAA	170641		/protein_id="BAB43135.1"
Qy	847	AAAGTCTGTATGGAAGAAATTAACCGCTGTGAATAATCAACGTTGTGATATTTAGAC	906		/db_xref="GI:13701843"
Ds	170640	GAAGATCAACATGCAAGAAATTTGAATGCAGTTAAATAATAATCGCGTGATTTCTGAC	170581		/translation="MSTNQFLFVIALILLITTSVIGVGRVMSRQRLKSMETLWOTI
Qy	907	CGTGACTTATGGCAAGATCAGTGGTTTAAATTTCTTCAAGAAATGCGCAAGAACTT	966		SPLAPTRPNSHFDEYKLYKEFESHSLVDVDDKTSWDLNMAIPIKMYNLTALGEMK
Ds	170580	CGTGATGTTGGCAAGATCTCGTGGCTTAAATTTCTTGAAGAAATGCGCTAAAGAACTT	170521		LYALRQMLSTINKSLSLFNDAFRKNVTYHIALIGKTYPTFPDQITPVKQNLIL
Qy	967	GTTGAATTTATCTAAGAAAGATAGTAAAAAAGAT	999		FLCPFLVIFGAVIFINSQVIGLLFLMCLFNIISATLTKRTYEDDLKSIIPMLDVLF
Ds	170520	GTTGAATTTATCAAAAAGAACAAAGTAAAGT	170488		QYTIKSIKHAQPOEVNFQKPTARHLTSLVAEVDNDEIGAMVILKVLKIPMLDVLF
RESULT 4					HSIQSVYTHMNLKNCDFYIAELDNHVALAMVRYRTLECYTEPOIDDTNDGDIVFSELT
AP003136/c					HPLADAVANDFSLSONILLTGSNASGKSTPMKSIANIILASAIQTVTASKEVYOPG
LOCUS	307750 bp	DNA	linear	BCT 11-JAN-2003	IVFTSMANADDVLSGDSYFMAELKSKIRIVEIPDNQKIYCFIDIFKGTNTTERLAAS
DEFINITION	Staphylococcus aureus subsp. aureus N315 genomic DNA, complete genome, section 8/10.				ESVLSFLHEKSNFRVIAATHDIELAELLKQRYENHFNEVIENNHNHFDYKIKPGKAN
ACCESSION	AP003136	BA000018			TRNAIBLKITSFPKATYIRAKDNVNSG"
VERSION	AP003136.2	GI:14349228			complement (2222..3247)
KEYWORDS	Staphylococcus aureus subsp. aureus N315				/gene="SA1854"
SOURCE	Staphylococcus aureus subsp. aureus N315				complement (2222..3247)
ORGANISM	Staphylococcus aureus subsp. aureus N315				/gene="SA1854"
REFERENCE	Bacteria; Firmicutes; Bacillales; Staphylococcus.				/note="ORFID:SA1854
AUTHORS	1 Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsunaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.				hypothetical protein, similar to O-sialoglycoprotein
TITLE	Whole genome sequencing of methicillin-resistant Staphylococcus aureus				endopeptidase"
JOURNAL	Lancet 357 (9264), 1225-1240 (2001)				/codon_start=1
MEDLINE	21311952				/protein_id="BAB43136.1"
PUBMED	11418146				/db_xref="GI:13701844"
REFERENCE	2 (bases 1 to 307750)				/translation="MDQSQEQLNIREMTKEDVPQVDFLRRSFNDSWTTDAFYHEI
AUTHORS	Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.				BQNNPARYFVLEPEQIIGYGLMWIVDQAQITTVADIDNRYGVLGQMLLYGINYA
TITLE	Direct Submission				SHTCDVMSLEVRVNNKVAQHVYENLFGYGGKRNKYGGEDAMVWVNLND"
JOURNAL	Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nite.go.jp, URL: http://www.bio.nite.go.jp/)				complement (3677..4366)
COMMENT	Tel: 81-3-3481-1933, Fax: 81-3-3481-8424				/gene="SA1856"
FEATURES	On Jun 12, 2001 this sequence version replaced gi:13701842.				complement (3677..4366)
source	Location/Qualifiers				/note="ORFID:SA1856
	1..307750				hypothetical protein, similar to glycoprotein
	/organism="Staphylococcus aureus subsp. aureus N315"				endopeptidase"
	/mol_type="genomic DNA"				/codon_start=1
	/strain="N315"				/transl_table=1

gene	LFGITACSSSIHFIN"	CDS	complement (7769..8500)	Query Match	56.4%; Score 569; DB 1; Length 347650;
	4617..4925			Best Local Similarity	73.3%; Pred. No. 8.3e-78;
	/locus_tag="SAV1941"			Matches	728; Conservative 0; Mismatches 265; Indels 0; Gaps 0;
	4617..4925				
	/locus_tag="SAV1941"				
	/note="Bacteriophage phi Sa 3mu"				
	/codon_start=1				
	/transl_table=11				
	/product="hypothetical protein"				
	/protein_id="BAB58103.2"				
	/db_xref="GI:46395540"				
	/translation="MKNTNINILLVINSLTHFLTLNTSFPNNSASDFCFIIGAIF				
	FLIGFVAIYMKRATYWLNLILFTNIFYFLHFCVLLLLKVIQFKLFIYEGCVLLET				
	"				
gene	complement (5216..5566)				
	/locus_tag="SAV1942"				
	complement (5216..5566)				
	/locus_tag="SAV1942"				
	/note="Bacteriophage phi Sa 3mu"				
	/codon_start=1				
	/transl_table=11				
	/product="hypothetical protein"				
	/protein_id="BAB58104.1"				
	/db_xref="GI:14247715"				
	/translation="MKIRKSILAGTLAIVLASPLVTNLDKNEAOASTSLPTSNEYONE				
	KLANELKSLLDELNVNELATGSLNTYYKRTIKISQKAMVALKSKDFKMWSEAKYQLQ				
	KIYNEIDALKSKY"				
gene	complement (6077..6373)				
	/gene="truncated-lytA"				
	complement (6077..6373)				
	/locus_tag="SAV1943"				
	/gene="truncated-lytA"				
	/locus_tag="SAV1943"				
	/note="Bacteriophage phi Sa 3mu"				
	/codon_start=1				
	/transl_table=11				
	/product="truncated amidase"				
	/protein_id="BAB58105.2"				
	/db_xref="GI:46395541"				
	/translation="MPAGYTLDKNNVPYKKGTYVTVANVKNVNRDGYNSNRITGV				
	LPNNAITKYDGAICYNGYRWITYIANNGQRCYIATGEVDKAGNRISSEGNFSAL"				
	6461..6616				
gene	/gene="truncated-lytA"				
	/locus_tag="SAV1943.1"				
	6461..6616				
	/gene="truncated-lytA"				
	/locus_tag="SAV1943.1"				
	/note="Bacteriophage phi Sa 3mu"				
	/codon_start=1				
	/transl_table=11				
	/product="truncated amidase"				
	/protein_id="BAD16608.1"				
	/db_xref="GI:46395542"				
	/translation="MRCSDCNISMLLRVMYRVKKGKDINCRISVRLGLHGLTGIVHVS				
	YIHLIIT"				
gene	complement (7063..7554)				
	/gene="sak"				
	/locus_tag="SAV1944"				
	complement (7063..7554)				
	/gene="sak"				
	/locus_tag="SAV1944"				
	/note="Bacteriophage phi Sa 3mu"				
	/codon_start=1				
	/transl_table=11				
	/product="staphylokinase precursor"				
	/protein_id="BAB58106.1"				
	/db_xref="GI:14247717"				
	/translation="MKRSILFLTVLLLPSPSSITNEVSASSPDKYKKGDDASY				
	FETPGYLMVNVTVGDKGNELLSPHYVEFPKPGPTLTKKIEIYVEWALDATAIKE				
	FRVVELDPKSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHKINPGFNLTIKVVI				
	EKK"				
gene	complement (7769..8500)				
	/locus_tag="SAV1945"				

RESULT 6
BX571856_23/c
WPCOMMENT

Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1110000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000
BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619

Continuation (24 of 29) of BX571856 from base 2300001 (BX571856 Staphylococcus aureus su

Query Match 56.3%; Score 567.4; DB 1; Length 110000;
Best Local Similarity 73.2%; Pred. No. 2e-77;
Matches 727; Conservative 0; Mismatches 266; Indels 0; Gaps 0;

Qy	7	GAATCAGTGAGAGGTTTAAATTTTAAAGTGTAAATGGCTTATGTTTGTGTTTAAATGCA	66
Db	50327	GGAACGATGAGAGGTTTAAATTTTAAAGTGTAAATGGCTTATGTTTGTGTTTAAATGCA	50268
Qy	67	ACTGCAGCATGTGGAATAATAGTTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA	126
Db	50267	GTTCGAGCTTGTGTTAATACGGATTAATTCAGTAAAGAGATCACTAAAGATACT	50208
Qy	127	GTGTAATCAAGCAGCAAGAGGTTACTACGAAAGTACCTAAACACCCCTAAACGGTGTGTT	186
Db	50207	ATTTTCGGTAAAGATGAAATGGTACAGTAAAGTACCTAAAGATGCAAAACGGTATCGTT	50148
Qy	187	GTTCTTGATGATTCATTTGTTGATCGGTTAGTTGCTTTAGATGTTTAAACCTGTTGGGATA	246
Db	50147	GTAATTAGATGATCTATTTGCGATGCAATGACGATTTAGACGTTTAAACCGATTTGGTATT	50088
Qy	247	CGGGATGATAACAAAAAATCGTATTATTAACCATTAAGAGATATAAATTTGGAATAAC	306
Db	50087	GCTGATGATGTTAGAAAAAAGCTATCAATTAACCGCTTAGAAAAAATTTCCGGATTAT	50028
Qy	307	ACTTCTGTAGGAACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAACTTAAACCAAGAT	366
Db	50027	ACTTCTGTAGGTACAGTAAACAGCAAACTTAGAGGAAATTAGTAAATTAACACCGGAT	49968
Qy	367	TTAATATTGCTGATAAATAATAGACAAAGGATTTATTAAGACTTAATAAATAAATTCCT	426
Db	49967	TTAATTATGCTGTATGACGATAGATAAAGGATTTAATAAAGAAATTTAAACAAATTGCA	49908
Qy	427	CTACGATTGAACCTGAAAGTTTCATGAGAGATTAATAATGAAAAATTTGATGCTTTTAAA	486
Db	49907	CCAACATTATCATTAAGAGTTTGTGTTGCTACTACAGCAAAATATCGATGCTTTCAAG	49848

Qy	487	ACAAATTTCAAAAGCTTTAGGTAAAGAAAGAGGTAATAAAACGCTTAGAAGAACACGAT	546
Db	49847	ACAAATTCGAAAGCTTTAGATATAAAGAAAAAGAGCGAAAAACGCTCTGTGTAACATGAT	49788
Qy	547	AAGAAAAATTGAAGAAATATAAAAAAGAAATAACTATGATATAAAATCAAAAGGTATTGCT	606
Db	49787	AAATTAATCAATAAGTATAAGATGAAATTAATTTGATAGAAATCAAAAAGTGTCTTCCA	49728
Qy	607	GCAGTAGCTGCTAAATCAGGTTTCTGCTCATCCAAAGCAACTCTTATTTGTTGGTCAATTC	666
Db	49727	GCAGTTGTTGGAAGAGCTGGTTTATTAGCACATCCAAACTATTCTGATGTTGGACAATTT	49668
Qy	667	CTAAGTCAACTAGGTTTTAAAGAAAGCAATTAAGTATGATGTTTACTTAAAGGTTTAAAGTAAG	726
Db	49667	TTAAACGAACCTTGGATTTAAAAATGCAATTAAGTATGATGTAACAAAAAGGTTTAAAGTAAA	49608
Qy	727	TATCTTAAAGGACCTTACTTACAAAATGAACACTCAAACTTTTATCTCAAGTCAATCTCGAG	786
Db	49607	TACTTGAAGGACCTTACTTACAAATTAAGTATGATGAACTTTAGCTGACTTAAATCTCTGAA	49548
Qy	787	CGTATGTTTCAATATGACAAAACAAAAGCAAGTTTCTTAACGAACCTTCACTAAAAAGAACTAGAA	846
Db	49547	CGCATGATTAATTAATGACAGATAATGCTAAAAAAGATTTCTGCTGAATTCAGAAAGTTACAA	49488
Qy	847	AAAGATCTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGTTGTGATTTTATAGAC	906
Db	49487	GAAGATGCAACATGGAAGAAAAATTTGAATGCACTTAAAAAATTAATCGTTGTGATTTGTGAC	49428
Qy	907	CGTGACTTTATGGCAAGATCACGTTGTTTAAATTTCTCAGAAAGAAATGGCAAAAGAACTT	966
Db	49427	CGTGATGTTTGGCAAGATCTCGTGGCTTAATTTCTTCTGAAAGAAATGGCTTAAAGAACTT	49368
Qy	967	GTTGAATTTATCTAAGAAAGATAGTAAAAAAGAT	999
Db	49367	GTTGAATTTATCAAAAAAAGAAACAAAGATAAGT	49335

RESULT 7
AX622368

LOCUS AX622368 1044 bp DNA linear PAT 20-FEB-2003
DEFINITION Sequence 5331 from Patent WO02094868.

ACCESSION AX622368

VERSION AX622368.1 GI:28450503

KEYWORDS

SOURCE Staphylococcus aureus

ORGANISM Staphylococcus aureus

REFERENCE 1 Masignani, V.C., Mora, M.C. and Scarselli, M.C.

AUTHORS Staphylococcus aureus proteins and nucleic acids

TITLE Patent: WO 02094868-A 5331 28-NOV-2002;

JOURNAL Chiron Spa (IT)

FEATURES Location/Qualifiers

source 1..1044

/organism="Staphylococcus aureus"

/mol_type="unassigned DNA"

/db_xref="taxon:1280"

ORIGIN

Query Match	55.3%;	Score 557.4;	DB 6;	Length 1044;
Best Local Similarity	73.5%;	Pred. No. 2.1e-75;		
Matches 711;	Conservative 0;	Mismatches 256;	Indels 0;	Gaps 0;
Qy	20	GTGTTAAAAATTTTAAAGTGTAAATTTGGCTTATTGTTGTTTGTGTTTAAATTTGCAACTGCAATGTG	79	
Db	71	GGTCTAAACCTTTTAGTATATTTGGGATTAATAGTTGCTTACTTTTAGTTGCAAGTGTG	130	
Qy	80	GAAATAATAGTTCAGTAACTCAAGTAAAGAGTATCAAAAAGATGGAGTTGAAATCAAGC	139	
Db	131	GTAATACGGATAATTTCAAGTAAAGAAATCATCAACTAAAGATACTATTTTCGGTAAAG	190	
Qy	140	ACGAAGAAGGTACTACGAAAGTACTTAAACACCCTTAAACGTTGTTGTTCTTCTTGAGTATT	199	

```

191 ATGAAATCGTACAGTAAAGTACCTAAAGATGCAAAAGCTATCGTTGTTAGAGTACT 250
200 CATTTGTTGATCGGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACA 259
251 CATTTGCGATGATAGTACAGATTTAGACGTTTAAACCGATTTGGTATGCTGATGATGTA 310
260 AAAAAATCGTATTATTAACCAATTAAGAGATTAATAATTGGAATAATACACTCTGTGAGAA 319
311 AGAAAAACGATATCATTAACACAGTTAGAGAAAAAATTTGGGATTTACTTCTGTAGTA 370
320 CACGTAAGCAACCTAACTTTAGAGAAATCAGTAAACTTTAAACCCAGATTTAAATTTGCTG 379
371 CACGTAACAGCCAACTTTAGAGAAATTTAGTAAATTTAAACCGGATTTAAATTTGCTG 430
380 ATAAATAGACACAAAGATTTATTAAGACATTTAAATTAATTTGCTCTAGATGTAAC 439
431 ATAGCAGTATAGATTAAGGTTATTAATAAGAAATTTAAACCAAAATTTGCACCAATTTATCAT 490
440 TGAAGATTTTCGATGAGATTTATTAAGAAATTTAGTCTTTTAAACCAATTTTCAAAAG 499
491 TAAAGATTTTCGATGAGATTTATTAAGAAATTTAGTCTTTTAAACCAATTTTCAAAAG 550
500 CTTTGTAGTAAGAAAGATTTAAACCGCTTTAGAGAAACACGATTAAGAAATTTGAG 559
551 CTTTAAATTAAGAAAGAAAGCGGAAACCGTCTTCTGAGCATGATTAATTAATCAATA 610
560 AATATAAAGAAATTAATCTGATGATTAAGAAATTAAGAAATTTGCTGAGTATGCTGTA 619
611 AGTATAAGATGAAATTTAAATTTGATAGAAATTAAGAAATTTGCTGAGCATGATTTA 670
620 AATCAGTTTGTCTGCTCATCAAGCAACTTTATGCTGCTCAATTTCCCTAAGTCAACTAG 679
671 AAGCTGGTTTATTTAGCACATCCAACTATTTATATGTTGGAACAATTTTAAACGAACTAG 730
680 GTTTTAAAGAAAGATTTAAGTATGATGTTACTAAAGGTTTAAAGTATGTTTAAAGGAC 739
731 GATTTAAATGATTTAAGTATGATGTTAAACAAAGGTTTAAAGTATGTTTAAAGGAC 790
740 CTTTACTTAACAATGAAACACTGAACTTTTATCTCAAGTGAATCTCTGAGCGTATGTTCA 799
791 CTTTACTTAACAATGAAACACTGAACTTTTATCTGATGTTTAAATTTCCAGAGCGTATGAT 850
800 TGAACAACAAGCAAGTTCTTAACGAACCTTCACTAAAGAACTAGAGAAAGATCTCTGTAT 859
851 TGACAGATCATGTAAGAAAGATTTCTGCTGAATTCAGAGATTTCAAGAAAGATGCAACAT 910
860 GGAAGAAATTAACCGCTGTGAAAAATCAACGCTGTTGATATTTTAGACCGTGAATTTATGG 919
911 GGAAGAAATTTGAATGATGATTTAAATAATCGCTGATATTTGACCGGATGTTTGGG 970
920 CAAGATCAGCTGTTTAAATTTCTTCAAGAAATTTGCAAGAAATTTGTTGAAATTTATCTA 979
971 CAAGATCTCGTGGCTTAAATTTCTTCAAGAAATTTGCAAGAAATTTGTTGAAATTTATCA 1030
980 AGAAGA 986
1031 AAAAAA 1037

```

```

RESULT 8
A64853/c LOCUS 2247 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 10 from Patent W09731114.
ACCESSION A64853
VERSION A64853.1 GI:4530844
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
1 Burnham, M.K. and Hodgson, J.E.
AUTHORS POLYNUCLEOTIDES AND AMINOACID SEQUENCES FROM STAPHYLOCOCCUS AUREUS
TITLE

```

JOURNAL Patent: WO 9731114-A 10 28-AUG-1997;
SMITHKLINE BEECHAM PLC (GB)
Location/Qualifiers

FEATURES
source 1..2247
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

```

Query Match 33.4%; Score 336.2; DB 6; Length 2247;
Best Local Similarity 70.6%; Pred. No. 6.8e-42;
Matches 580; Conservative 0; Mismatches 229; Indels 13; Gaps 10;
QY 1 GGAGTGAATCAGTCAGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGTTGTTT 60
DB 831 GGAGTGGGACNGATGAGAGGTTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCTTTA 772
QY 61 ATTGCAACTGAGCAGTATGGAATAA---ATTAGTTCAAGTAACTCAAGTAAAGAGTCATCA 117
DB 771 CTTTGTAGTTGAGCTTTGTGTAATACGGATAATTTCAAGTAAAGAGAG--AATCATCAACTA 713
QY 118 AAAGATGGAGTTGAAATCAAGCAGCAAGAGGTACTACGAAAGTACCTTAAACACCCCTAAA 177
DB 712 AAGATACATTTTCGGGTAAAGATGAAATTTGGTACAGTAAAGATACCTTAAAGATGCACBA 653
QY 178 CGTGTGTTGTTCTTGTAGTATTTCAATTTG--TTGATGCGTTAGTTGCT--TTTAGATGTTAAA 234
DB 652 CGTATCGTTGTTATTTAGAGTACTCAATTTTCAAGGTTTGCATTTANCAGCATATAGCCGTTANA 593
QY 235 CC-TGTTGGGATAGCGGATGATACAAAAAATCGTATTTATTAACCACTTAAAGATATAA 293
DB 592 CCAAGTTGGTATTTGCTGATGATGTTGTAAGAAAAAACGTTATCA--TAAACCCAGTAAAGAA 535
QY 294 AATTGGAATAATACCTCTGTAGGAAACAGTAAAGCAACTTACTTAGAAGAAATCAGTAA 353
DB 534 AATTGGGGGTATATCTTCTGTAGGTACACGTAA--CAAGCNAACTTTAAGAGANATTTAGTAA 476
QY 354 ACTTAAAAACAGATTTAAATTTATTTGCTGTAATAATAATAGACACAAAGGATTTTATAAAGACTT 413
DB 475 ATTAANACCGGATTTAAATTTATCNCCTGATAGCAGTAGACATAAGGGTATTTAATAAAGAAAT 416
QY 414 AATAAATTTGCTCTCTAGTGAATGAAAGTTTCCATGAGAGATTTATATGAAATAT 473
DB 415 AGCCCAACATTTGCACC--ACATTTATCATTTAAAGAGTTTGTATGGAGACTACACCCCAATAT 357
QY 474 TGATGCTTTTAA--AACATTTTCAAGGCTTTTAGGTAAAGAAAGAAAGAAAGTAAACCGCT 532
DB 356 TAAATTCGTTCAACAAATTTGCTAAAGCTTTTAAATAAGANAAAGAGCGGNNAACGCTC 297
QY 533 TAGAAGAAACACGATTAAGAAATTTGAAGAATATAAAAAAGAAATAACTATGATAAAAAATC 592
DB 296 TTGCTGAGCATGATTAATTTAATCAATAGTATAAAGATGAATTTAAATTTGATAGAAATC 237
QY 593 AAAAGATTTGCTGAGTATGCTGTAATTAATCAAGTTTGTCTGCTCATCAAGCAACTCTT 652
DB 236 ACAAGATGCTTCCAGCAGTAGTTGCTAAAGCTGTTTATTAGCAGTATCCACACTATTCAT 177
QY 653 ATGTTGGTCAATTTCTTAAGTCAACTAGGTTTAAAGAGCAATTTAAGTATGATGTTTACTA 712
DB 176 ATGTTGGCAATTTTAAACGAGCTAGGATTTTAAAGATTTGATTTAGTATGATGATGATGAC 117
QY 713 AAGGTTTAAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTC 772
DB 116 AAGGTTTAAAGTAAATTTTAAAGGACCTTACTTACAAATGAGCACTGAAACATTTAGCTG 57
QY 773 AAGTGAATCCTGAGCGTATGTTTCAATAATGACAAACAAAGCAA 814
DB 56 ATTTAAATCCAGAGCGTATGATCAATTTATGACAGATCATGCTA 15

```

RESULT 9
AP003132
LOCUS

298050 bp DNA linear BCT 11-JAN-2003


```
/gene="SA0799"  
/note="ORFID:SA0799"  
hypothetical protein, similar to NADH dehydrogenase"  
/codon_start=1  
/transl_table=11  
/protein_id="BAB42038.1"  
/db_xref="GI:13700742"  
/translation="MKNLVLLGGYGNRIMSRILTTSLPQDYTVTLVDRMPFHLKLP  
EYVALAGTKSPDKDVRMPNPNVYVGEINDIDLDQAIIVSGNSKIDDELIGL  
GEDCKHNVPAEETHYSIQTLSKARDTFHSISELPEGAKGIVGAGLSGLELASELR  
ERSDLEIYLDRGRILRNPEFKLSKVAKFWFAGNNVTVPNSINKVFGKINYCD  
BKPDIDLVTWVTFAGIQPEVVRNLPIDINSGRVIVNQHVPTRYRNVYVGGDADLPH  
APSAQLAIEVQGGIQIADVLKQWLNELPLDKMPELVKIVGSLGDKQGFAYIMDRVT  
GLASTLKSGLVLYKYHNG"  
6911..7147  
/gene="SA0800"  
6911..7147  
/gene="SA0800"  
/note="ORFID:SA0800"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="BAB42039.1"  
/db_xref="GI:13700743"  
/translation="MNPVIFCLSNMAKGGYVFNQLENDPDPDVLVYGLTHCGICS  
AGLYALVNGDIVEGDSPBELLQNIYAHKETWIF"  
7160..7519  
/gene="SA0801"  
7160..7519  
/gene="SA0801"  
/note="ORFID:SA0801"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="BAB42040.1"  
/db_xref="GI:13700744"  
/translation="MPTVILTAAAYEVRDMLKANEMPDYILKIVNGGGCTGLTYGM  
GAEAPGENDEVLEFGLKVLVDKEDAPVLNGTTIDFKQSLMGGGFOIDNPNAIASCG  
CSSFTKAVKVPENC"  
7973..9181  
/gene="SA0802"  
7973..9181  
/gene="SA0802"  
/note="ORFID:SA0802"  
hypothetical protein, similar to NADH dehydrogenase"  
/codon_start=1  
/transl_table=11  
/protein_id="BAB42041.1"  
/db_xref="GI:13700745"  
/translation="MAQDRKKVLVAGYAGLQTVTKLOKAIISTEEAEITLKNKNEYH  
YEATWLHEASAGLNTIEDVLVPVESVLKDKVNFQAEVTKIDRDAKKVETNQGIIYDF  
DILVALGFVSFTFEGMKDHPAFQENIVITELSRHEDFPNTAAKSKEDDNDLS  
ILVGGAGFTGVFELGELTDRIPELSKYGVDQNKVKITCEAAAPKPLPMFSEELVNAH
```

```
Query Match 26.7%; Score 269.2; DB 1; Length 298050;  
Best Local Similarity 55.4%; Pred. No. 2.9e-32;  
Matches 545; Conservative 0; Mismatches 433; Indels 6; Gaps 1;  
  
QY 17 GAGGTTTAAAAATTTTAAAGTGTAATTCGGCTTATTGTTGTTTAAATTCGCAACTGCAGCAT 76  
DB 113840 GAGGTAAATATGTCATAGGAATATCGTTAAATTAGTTGTTTATGCTAATCTTAGTTG 113899  
  
QY 77 GTGGAATAATAGTTCAAGTAACTCAAGTAAAGTCATCAAAAGATGGAGTTGAATCA 136  
DB 113900 TAGCAGTAGCGGGTGTGGTCAAAAAGATCTGAAGAGAAAACCTGAAATGACGACAATA 113959  
  
QY 137 AGCAGAGAGGACTACGAAAGTACCTAAACACCCCTAAACGCTGTTGTTGTTCTTGAGT 196  
DB 113960 AAGATGAATTAGGAACTCGAAAATAATTAAGAAAATCTCTAAACGTTGTTGTTATTAGAT 114019  
  
QY 197 ATTCATTTGTTGATCGCTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256  
DB 114020 ATAGTTTGTGATTTATTAGCAGCATTAGATATGAACCTGTTGTTGTTGATGATGATG 114079
```

```
QY 257 ACAAATAAAATCGTATTATTATAAACCATTTAAGAGATAAAATTTGAAATAATACACTTCTGTAG 316  
DB 114080 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTTGGGGCATATGAATCGGTTG 114139  
  
QY 317 GAACACGTAACCAACCTTAACCTTAGAAGAAATCAGTAACCTTTAAACCAAGATTTAAATTATG 376  
DB 114140 GATCTAGACCCCAACCGGATATGGAAGTGAAGAATTAATAACCGGATTTGATCATG 114199  
  
QY 377 CTGATAATAATAGACACAAAGGTATTTATAAGAGCTTTAAATAAAATTCCTCTCAGATTG 436  
DB 114200 CAGATGTTAGCAGACATGAAGAATCAATCAGAAATTCGACCAAAATTCCTCCGACAAATCA 114259  
  
QY 437 AACTGAAAGTTTCGATCGCAGATTATATGAATAATATTGATGCTTTTAAAAACAATTTCAA 496  
DB 114260 TGTAGTTAGCGGTACGGGAGATTATATGCAAAATATTGAAGCATTTTAAAAACAGTCGCTA 114319  
  
QY 497 AAGCTTTAGGTAAAGCAAGAGGTAAGAAACCGTTAGAGAGACACCATAGAAATTTG 556  
DB 114320 AAGCAGTAGGCAAGAGAGAGCGGCTCGGAAGAGCGCTCGGAAGAGCATGATATAAATATTAG 114379  
  
QY 557 AAGATATAAAAAAGAAATAACTATGATGATAAAAAATCAAAAGGATTTGCTTCAGTAGCTG 616  
DB 114380 CGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAATCTGCAATTCGATTCGTAATCT 114439  
  
QY 617 CTAAATCAGGTTTGTCTCTCATCAAGCAACTCTTATGTTGGTCAATTCCTTAAGTCAAC 676  
DB 114440 CAAGAGCAGGTATGTTTATTATGAAGATACATTTTATGGCAAAATCTTAAATTAATA 114499  
  
QY 677 TAGGTTTAAAGAACATTAAGTATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736  
DB 114500 TGGGTATTTCAACCTGAGTCAAAAGACAAACATCGCATGTTGGTGAACCAAGGGTG 114559  
  
QY 737 GACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796  
DB 114560 GTCTTATATATTTTAAATGAAGAACTTGCCCAATATCATCCAAAAGTTATGATTT 114619  
  
QY 797 TAATGACAAAACAAAGCAAGTTCTTAACGAACTTCACTAAAGAACTAGAAAAAGATCCCTG 856  
DB 114620 TAGCCACTGACGGAAGAAACGGACAAAAA-----TAGAACGAAATTCATTGATCCTGCAG 114673  
  
QY 857 TATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTAGCCGCTGACTTAT 916  
DB 114674 TTTGGAATTCATTAAGAGCTGTGAAGATAACAAAGTTTATGACGTTGACCGGAAATAAGT 114733  
  
QY 917 GGGCAAGATCAGTGTGTTTAAATTTCTCAGAGAAATGCGAAAGAACTTTGTTGAATTAT 976  
DB 114734 GGTGAAATCAAGGGGGATTATCGCAAGTGAAGTATGCGCAGAGATTTAGAAAAAATTTG 114793  
  
QY 977 CTAAGAAAGATAGTAAATAAGATA 1000  
DB 114794 CAGAAAAGCAAAATATAAATAACA 114817
```

```
RESULT 10  
AP003361  
LOCUS  
DEFINITION  
Staphylococcus aureus subsp. aureus Mu50 DNA, linear BCT 17-APR-2004  
section 4/9.  
ACCESSION  
AP003361 BA000017  
VERSION  
AP003361.3 GI:46395333  
KEYWORDS  
Staphylococcus aureus subsp. aureus Mu50  
Staphylococcus aureus subsp. aureus Mu50  
Bacteria; Firmicutes; Bacillales; Staphylococcus.  
1
```

```
REFERENCE  
AUTHORS  
Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,  
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,  
Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,  
Mizutani, U., Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C.,  
Sekimizu, K., Hirakawa, H., Kihara, S., Goto, S., Yabuzaki, J.,  
Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,  
Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.
```


Db 128 AAGTGAATAGGAAGTGAAGAAATTAAGAAAAATCCTAAGCGTGTGTTGTTATAGAAAT 187
Qy 197 ATTCAATTTGTTGATGCGTTAGTGTCTTGTAGATGTTTAAACCTGTTGGGATAGCGGATGATA 256
Db 188 ATAGTTTTGCTGATTTATTTAGCAGCATTTAGATATGAAACCTGTTGGTATTGCGAGATGATG 247
Qy 257 ACAAAAAAATCGTATTTATTAACCATTTAAGAGATAAAATTTGGAATAATACACTTCTGTAG 316
Db 248 GCAGCACTAAATAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGGTTG 307
Qy 317 GAACACGCTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTTAAACCGAGATTAAATTTATG 376
Db 308 GATCTAGACCGAACCGGATATGGAAGTGATAAGTAAATTTAAACCGGATTGATCATG 367
Qy 377 CTGATAATAATAGACACAAAGGTATTTATAAGACTTTAAATTTAAATTTCTCTCTACGATG 436
Db 368 CAGATGTTAGCAGACATTAAGAAATCAAAATCAGAAATTCAGCAAAATTCCTCGACAATCA 427
Qy 437 AACTGAAAGTTTCGATCGAGATTATATGAAATTAATGAAATTTGATGCTTTTAAACAATTTCAA 496
Db 428 TGTAGTTAGCGGTACGGGAGATTATATGCAAAATATTGAAGCAATTTTAAACAGTCGCTA 487
Qy 497 AAGCTTTAGGTAAAGAAAGGTAAAGGTAAAGAAAGCTTAGAAGAACACAGATTAAGAAATTTG 556
Db 488 AAGCGTAGGCAAGAAAGAAAGCGGAGAGCGTCTGGAAGAGCATGATTAATAATTTAG 547
Qy 557 AAGATATAAAAAAGAAATAACTATGGAATAAAAAATCAAAAGGTATTCGCTGCAGTAGCTG 616
Db 548 CGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAATCTGCAATTTGCAATTCGGTATCT 607
Qy 617 CTAATCAGGTTTGTCTGCTCATCAAGCAACTCTTATGTTGGTCAATTCCTAGTCAAC 676
Db 608 CAAGAGCAGGTATGTTTATTAATTAATGAAGATACATTTTATGCGGCAATTCCTTAATAAA 667
Qy 677 TAGGTTTAAAGAAAGCATTTAAGTATGATGTTTACTTAAAGGTTTAAAGTATCTTAAAG 736
Db 668 TGGGTATTCACCTGAGTCAAAAGAACAAACTACCGATGTTGGTGAACGCAAGGTTG 727
Qy 737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
Db 728 GTCCTTATATATTTTAAATTAATGAAGAACTTGCCAAATATCAATCCAAAGTTTATGATTT 787
Qy 797 TAATGACAAACAAAGCAGTTCTAACGAACCTTCACTTAAAGAACTAGAAAGAAAGTCTG 856
Db 788 TAGCCACTGACGGAAAAACGGACAAAA-----TAGAACGAAATTCATTGATCCTGCAG 841
Qy 857 TATGGAAGAAATTAACCGCTGTGAAATCAACGTTGTTGATATTTTAGACCGTGACTTAT 916
Db 842 TTTGGAATCATTTAAAGCTGTGAAGATACAAAGTTTATGACGTTGACCGAAATAAGT 901
Qy 917 GGGCAAGATCAGTGGTTTAAATTTCTTCAGAGAAATGGCAAAAGAACTTGTGTAATTTAT 976
Db 902 GGTGAAATCAAGGGGATTTACGCAAGTGAAGTATGCGAGAGATTTAGAAAAAATTTG 961
Qy 977 CTAGAAAG 985
Db 962 CAGAAAAAG 970

RESULT 12
AX583924
LOCUS AX583924
DEFINITION Sequence 301 from Patent WO02059148.
ACCESSION AX583924
VERSION AX583924.1 GI:27655594
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Meinke, A., Nagy, E., von Ahsen, U., Klade, C., Henics, T., Zauner, W.,
Minh, D.B., Vytvytska, O., Etz, H., Dryla, A., Weichhart, T., Hafner, M.,

Tempelmaier B., Fraser, C.M. and Gilli, S.
A method for identification, isolation and production of antigens
to a specific pathogen
Patent: WO 02059148-A 301 01-AUG-2002;
Cistem Biotechnologies GmbH (AT)
Location/Qualifiers
1. 957
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"
ORIGIN
Query Match 26 14; Score 263.2; DB 6; Length 957;
Best Local Similarity 55.9%; Pred. No. 1e-30; Mismatches 408; Indels 6; Gaps 1;
Matches 524; Conservative 0;
Qy 48 ATTGTTGTTTAAATTCGCACTGCAGCATGTGGAAATAATAGTTTCAAGTAACTCAAGTAA 107
Db 21 ACTAGTTGTTGTTTCATGCTAATTTTGTAGTGTAGCGGTGTTGTTGCTCAAAAGATAC 80
Qy 108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCAAGAAAGGTTACTAGAAAGTACCTAA 167
Db 81 TGAAGAGAAAACTCAAAATGACGACATAAAAGATGAAATTTAGGAATCTGAAAAAATTAAGAA 140
Qy 168 ACACCTTAAAGCTGTTGTTGTTCTTGTAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
Db 141 AAATCCTTAAAGCTGTTGTTGTTATTTAGAAATATAGTTTTCGTTGATTTATTAGCAGATTAGA 200
Qy 228 TGTTTAAACCTGTTGGGATAGCGGATGATTAACAAAAAATCGTATTATTAAACCATTAAG 287
Db 201 TATGAACCTGTTGGTATTCAGATGATGGCAGCACTTAAAAATATACAAAGTCAGTAAG 260
Qy 288 AGATAAAATTTGAAAAATACACTTCTGTAGGAACACGTAAAGCAACCTAACTTAGAAGAAAT 347
Db 261 AGATAAGATTTGGGCATATGAATCGTTGGATCTAGACCAGCAACCGAATATGGAAGTGAT 320
Qy 348 CAGTAAACCTTAAACCAAGATTTAAATTTGCTGATAATAATAGACACAAAGGTTATTATAA 407
Db 321 AAGTAAATTTAAACCGGATTTGATTCATTTGCAGATGTTTAGCAGACATTAAGAAATTCAAATC 380
Qy 408 AGACTTAAATTAATAATTCCTCCTACGATTGAACTGAAAGTTTTCGATGGAGATTATAATGA 467
Db 381 AGAATTGAGCAAAATTCCTCGACAAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440
Qy 468 AAATATTGATGCTTTTAAAAACAATTTCAAAAGCTTTTAGGTTAAAGAAAGAGTAAAAA 527
Db 441 AAATATTGAAGCATTTTAAACAGTCGCTTAAAGCAGTAGGCAAGAGAAAGAAAGCGAGAA 500
Qy 528 ACGCTTAGAAGAACACCGATAGAAATTTGAAGATATAAAAAAGAAATAAATACTATGGATAA 587
Db 501 GCGTCTGGAAGACATGATTAATAATTAGCGGAGATTAGAAAGAAATTTGAACAGAGTAC 560
Qy 588 AAATCAAAAGGATTTGCTGCGATGCTGCTTAAATCAGGTTTTCGTTGCTCATCCAAGCAA 647
Db 561 GTTAAATCTGCAATTCGATTCGATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620
Qy 648 CTCCTATTGTTGGTCAATTCCTAAGTCAACTAGGTTTAAAGAAAGCATTTAAGTGTATGATGT 707
Db 621 TACATTTATGGGCAAAATTCCTTAATTAATAATGGGTATTCAACCTGAAAGTCACAAAAGCAA 680
Qy 708 TACTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTTACAAATGAACCTGAAACCTTT 767
Db 681 AACTACGCGATGTTGGTGAACGCAAGGGTGGTCTTATATATATTATTAATAATGAAGAACT 740
Qy 768 ATCTCAAGTGAATCCTGAGCGTATGTTTCAATAATGACAAACAAAGCAAGTCTTAAAGCAACC 827
Db 741 TGCCAAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797
Qy 828 TTCACTAAAAGAACTAGAAAAAGATTCCTGTATGGAAGAAATTTAAACGCTGTGAAAAATCA 887
Db 798 ---TAGAAGCAAAATTCATTGATCCTGCAAGTTTGGAAATCATTTAAAGAGCTGTGAAAGATAA 854
Qy 888 ACGTGTGATATTTTAGACCGTGACTTATATGGGCAAGATCACGTGTTTAAATTTCTTCAGA 947

PFYKFFIVQSTFERGFVLSVIVLLSLVLSVIRIFLQGFQPKGYDLNNKVDVK
YLTTIAIVAVVITVLYGLSADLYPMVKAGATFYNPSTVYKAVLGK"

gene

complement (9275..9616)

CDS

/gene="mnhc"

complement (9275..9616)

/gene="mnhc"

/note="ORFID:MW0832"

/codon_start=1

/transl_table=11

/product="Na+/H+ antiporter subunit"

/protein_id="BAB94697.1"

/db_xref="GI:2120399"

/translation="MEIIMIFVSGILTAISVYLSKSLIRIVMGTTLLTHANLFLI

TMGLKHGTVPDIYEANVKSYYDPIQALILTAIVAFATTAFFLVLAFTYKELGTDN

VESMKGVPEDD"

complement (9616..10044)

/gene="mnhb"

complement (9616..10044)

/gene="mnhb"

/note="ORFID:MW0833"

/codon_start=1

/transl_table=11

/product="Na+/H+ antiporter subunit"

/protein_id="BAB94698.1"

/db_xref="GI:2120400"

/translation="MNRQNDLILQFAAVIIFVMVFGSLFLAGHYTPGGFVGL

Query Match 26.1%; Score 262.8; DB 1; Length 290150;
Best Local Similarity 55.0%; Pred. No. 2.8e-31;
Matches 541; Conservative 0; Mismatches 437; Indels 6; Gaps 1;

Qy	17	GAGGTTTAAATTTTAAAGTGAATTTGGCTATTGTTGTTTAAATGCACTGCAGCAT	76
Db	107984	GAGGTAATGCTGAATGAATATCGTTAACTAGTTGTTTAACTATCTTAGTTG	108043
Qy	77	GTGGAATAATATCAAGTAACTCAAGTAAAGAGTCACTCAAAAGATGGAGTTGAATCA	136
Db	108044	TAGCAGTAGCGGTTGTGCTCAAAAGATACCTGAAGAGAAACTGAAATGACGACATAA	108103
Qy	137	AGCAGAGAAGGTACTAGAAAGTACTAAACACCTTAAACGTTGTTGTTCTTGAGT	196
Db	108104	AAAGTCAATTAGCACTGAAATTTAAGAAATCTTAAACGTTGTTGTTGTTAGAT	108163
Qy	197	ATTCAATTTGCTGCTAGTTGTTGTTAGATTTAAACCTCTGGATAGCGGATGATA	256
Db	108164	ATAGTTTCTGCTGATTTATAGCAGCAATAGATGAACCTGTTGTTGTCAGATGATG	108223
Qy	257	ACAAATAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAATAACACTTCTGTAG	316
Db	108224	GCAGCAGTAAATATAACAAAGTCACTGAAGATAAGATTGGGCATATGAATCGGTTG	108283
Qy	317	GAACAGTAAAGCACTACTTAGAAGAAATCAGTAAACCTTAAACGAGATTTAATTTG	376
Db	108284	GATCTAGACCGCAACCGAATATGGAAGTGAATAAGTAAATTAACCGGATTTGATCATTG	108343
Qy	377	CTGATAATTAAGCAAGGTATTATAAGACTTAAATAAATTTGCTCTAGCATTTG	436
Db	108344	CAGATTTAGCAGACATAAGAAATCAATCAGATTGAGCAAAATTTGCCGACAAATTA	108403
Qy	437	AACTGAAATGTTTCGATGAGATTTAATCAAAATATTGATGCTTTTAAACCAATTTCAA	496
Db	108404	TGTTAGTCAGTGTACGGGATTTACATGCAATATTGAGCAATTTAAACAGTCGCTA	108463
Qy	497	AGCTTTAGTAAAGAGAGAGGTAAACCGCTTAGAAGAACACAGATAGAAATTTG	556
Db	108464	AGCAGTTGCAAGAGAGAGAGCGGAGAGCGTCTGGAAGAGCATGATAAATATTAG	108523
Qy	557	AGAATAATAAAGAAATCACTATCGATAAAATCAAAAGGTATTGCTCGCAGTAGCTG	616
Db	108524	CGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAATCTCGATTTGCAATTCGGTATCT	108583
Qy	617	CTAAATCAGGTTTGTCTCATCCCAAGCACTCTTATGTTGTTGTTCAATTCCTTAAGTCAAC	676

Db	108584	CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAAATCTTACTTAAA	108643
Qy	677	TAGTTTTAAAGAACCAATTAAGTATGATGTTTACTAAAGGTTTAAAGTATCTTAAAG	736
Db	108644	TGGTATTTCAACCTGAAGTCACAAAAGACAAAATCTACGCATGTTGGTGAACGCAAGGTG	108703
Qy	737	GACCTTACTTACAAATGAACACTGAACCTTTATCTCAAGTGAATCTCTGAGCGTATGTTCA	796
Db	108704	GTCTTATATATATTTAAATAATGAAGACTTGCCTAATATCAATCCAAAAGTTATGATTT	108763
Qy	797	TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAGAACTAGAAAAGATCCTG	856
Db	108764	TAGCCATGACGGGAAACCGACAAAA-----TAGAACGAATTCATTGATCCTGCAG	108817
Qy	857	TATGAAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTAGACCGTCACTTAT	916
Db	108818	TTTGGAATCATTTAAAGCTGTGAAGATAACAAAGTTTATGACGTTTGACCGAAATAAGT	108877
Qy	917	GGGCAAGATCACGGTGTAAATTTCTTCAGAAAGAAATGGCAAAAGAACTTGTGCAATTAT	976
Db	108878	GGTTGAAATCAAGGGGATTTATCGCAAGTGAAGATATGGCAGAAAGATTTAGAAAAATTG	108937
Qy	977	CTAAGAAAGATAGTAAAAAGATA	1000
Db	108938	CAGAAAAGCAAAATAAAATACA	108961

RESULT 15

BX571856_10

WPCOMMENT

Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1110000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000
BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619

Continuation (11 of 29) of BX571856 from base 1000001 (BX571856 Staphylococcus aureus s

Query Match	25.2%;	Score 254.2;	DB 1;	Length 110000;	
Best Local Similarity	54.3%;	Pred. No. 7.2e-30;	Mismatches 448;	Indels 6;	Gaps 1;
Matches 539;	Conservative 0;				
Qy	8	AATCAGTCAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTTAAATTCGCA	67		
Db	54194	AAATGTTGGAGGTAATGTTGATAGGAATATCGTTAAATAGTTGTTGTTTATGCTAA	54253		
Qy	68	CTGCAGCATGTGGAATAATATAGTTCAAGTAAAGATCAATCAAGATGGAG	127		

```
Db 54254 TCCTAGTTGTAGCAGTACGCGGTTGTGTCACAAAGATACCTGAAGAGAAACTGAAATGA 54313
Qy 128 TTGAAATCAAGACGAAAGAGGTACTACGAAAGTACCTAAACACCCCTAAACGTTGTTG 187
Db 54314 CGACGATAAAGATGAAATAGGAATCTGAAATTAAGAAAATTCGAAAGATTTGTTG 54373
Qy 188 TTCTTGAGTATTCATTTGTTGATCGGTTAGTTGCTTTAGATGTTTAAACCTTTGGGATAG 247
Db 54374 TATTAGAATATAGTTTGTCTGCTGTTTATTAGCAGCATTTAGATATGAACCTTTGGTATTG 54433
Qy 248 CGGATGATAACAAAATAATCGTATTATTAAACCAATTAAAGAGATAAAATTCGAAAATACA 307
Db 54434 CAGATGATGTTAGTACTTAAATAATATAAAGTACGTAAGAGATAGATTGGGCATATG 54493
Qy 308 CTTCTGTAGGAACACGTAAGCAACCTAACTTTAGAGAAATCAGTAAACCTTTAAACAGATT 367
Db 54494 AATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGTAAAGTAAATTAATAACCGGATT 54553
Qy 368 TAATTTATGCTGATAATATAGACACAAAGGTATTATTAAAGACTTAAATTAATTTGCTC 427
Db 54554 TGNTTATTGTCAGATGTCAGCAGACATAAAGAAATCAAAATCAGAAATTTGAGCAAAATTTGCG 54613
Qy 428 CTACGATTGAACCTGAAAGTTTCGATGAGATTAATGAAATATTTGATGCTTTTAAAA 487
Db 54614 CGACATCATGTTAGTTAGCGGTACGGAGATTACATGCAATATTGAAGCATTTTAAAA 54673
Qy 488 CAATTTTCAAAAGCTTTAGTTAAGAGAAAGGTAAAGAAACGCTTTAGAGAAACACGATA 547
Db 54674 CAGTCGCTAAAGCAGTTGGCAAGAGAGAAAGCGGAAAGCGTCTGGAAGACATAATA 54733
Qy 548 AGAAATTTGAAGATAATAAAGAAATTAATGATGATAAATAATCAAAAGTATTGCTG 607
Db 54734 AAATATTAGCGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAATCTGCAATTTGCA 54793
Qy 608 CAGTAGCTGCTAAATCAGTTTGTCTCATCAAGCAACTCTTATGTTGTTCAATTC 667
Db 54794 TTGTTATCTCAAGCAGGTATGTTTATTAATATGAAGATACATTTATGGGACATTTCT 54853
Qy 668 TAAGTCAACTAGTTTAAAGAACGATTAAGTATGATGTTACTAAAGGTTTAAAGTAAAT 727
Db 54854 TAATTAATAATGGTATTCAACTGAAGTCACAAAGACAAAACCTGCGCATGTTGTTGAAC 54913
Qy 728 ATCTTAAGGACCTTACTTACAAATGACACTGAACTTTATCTCAAGTGAATCTGAGC 787
Db 54914 GCAAGGTTGGCCCTTATATTTTAAATATGAAGAGCTTGCCCAATATCAATCAAAAG 54973
Qy 788 GTATGTTTCATATGACAAACAAAGCAAGTCTCAACGAACTTCACTAAAGAACTTAGAAA 847
Db 54974 TTATGATTTTAGCTACGATCGAAATGGAATAACAGATTAATA-----TAGAACGAAATTCATTG 55027
Qy 848 AAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTAGACC 907
Db 55028 ATCTGCGAGTTTGGAAATCATTAAGCTGTGAAAGATAATAAAGTATATGATGTTGATC 55087
Qy 908 GTGACTTATGGCAAGATCAGCTGTTTAAATTTCTTCAAGAGAAATGCGAAAGAACTTG 967
Db 55088 GAAATAAGTGTGTTCAATCAAGAGGTATTTAGGCAAGTGAAGTATGCGAAGATTTAG 55147
Qy 968 TTGAATTATCTAAAGAAAGATAGTAAAAAGATA 1000
Db 55148 AAAAAATTCAGAAAAGCAAAATTAATAATACA 55180

RESULT 16
LOCUS AR354237
DEFINITION Sequence 355 from patent US 6593114.
ACCESSION AR354237
VERSION AR354237.1 GI:33760321
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
```

```
REFERENCE 1 (bases 1 to 668)
AUTHORS Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6593114-A 355 15-JUL-2003;
FEATURES
    source 1..668
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
    Query Match 22.1%; Score 222.6; DB 6; Length 668;
    Best Local Similarity 74.8%; Pred. No. 1.7e-24;
    Matches 279; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 627 TTTGCTTGCTCATCCAAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAA 686
Db 1 TTTATTAGCACATCCAAACTATTTCATATGTTGGACAATTTTAAACGAACTAGGATTTAA 60
Qy 687 AGAAGCATTAAGTCATGATGTTACTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTT 746
Db 61 AAATGCAATTAAGTGACGATGTAACAAAGGTTTAAAGTAAATATTTTGAAGGACCTTACTT 120
Qy 747 ACAATGCAACACTGAAACTTTTCTCAAGTGAATCCTGAGCGTATGTTTCAATATGACAAA 806
Db 121 ACAATTAGACACTGAAACATTTAGCTGATTTAAATCCAGAGGATGATCATATTATGACAGA 180
Qy 807 CAAAGCAAGTTTCTAACGAACCTTCACTAAAGAACTAGAAAAGATCCTGTATGGAAGAA 866
Db 181 TCATGCTAAAAAAGATTCTGCTGAATTTCAAGAGTTTCAAGAGATGCAACATCGAAAAA 240
Qy 867 ATTAAACGCTGTGAAAATCAACGTTGTGATATTTTAGACCGTCACTTATGGGCAAGTC 926
Db 241 GTTGAATGCGATTAATAAATAATATCGCGTGGATATTGTTGACCGTATGTTGGGCAAGATC 300
Qy 927 ACGTGGTTTAAATTTCTTCAAGAGAAATGGCAAGAAACTTGTGCAATTTATCTAAGAAAGA 986
Db 301 TCGTGGCTTAAATTTCTTCTGAGAAATGGCTAAAGAACTGTTGCAATTTATCAAAAAAGA 360
Qy 987 TAGTAAAAAAGAT 999
Db 361 ACAAAGTAAAGT 373

RESULT 17
LOCUS AR535793
DEFINITION Sequence 355 from patent US 6737248.
ACCESSION AR535793
VERSION AR535793.1 GI:53927010
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 668)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 355 18-MAY-2004;
FEATURES
    source 1..668
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
    Query Match 22.1%; Score 222.6; DB 6; Length 668;
    Best Local Similarity 74.8%; Pred. No. 1.7e-24;
    Matches 279; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 627 TTTGCTTGCTCATCCAAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAA 686
Db 1 TTTATTAGCACATCCAAACTATTTCATATGTTGGACAATTTTAAACGAACTAGGATTTAA 60
```

QY 687 AGAAGCAATTAAGTATGATGTTTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTT 746
 DB |||||
 61 AATGCAATTAAGTATGATGTTTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTTACTT 120
 QY 747 ACAATGAACACCTGAAACCTTTATCTCAAGTGAATCTGAGCGTATGTTTCAATGACAAA 806
 DB |||||
 121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGA 180
 QY 807 CAAGCAAGTTCTTAACGACCTTCACTAAAGCACTAGAAAAAGATCTGTATGGAAGAA 866
 DB |||||
 181 TCATGCTAAAAAGATTCTGCTGAATCAAGAGTTTCAAGAGATGCAATGGAAGAAA 240
 QY 867 ATTAAACGCTGTGAAAGTCAACGTTGATATTTTAGACCGTGAATTTAGGCGCAAGATC 926
 DB |||||
 241 GTTGAATGACGTTAAATAAATCGCTGATATTTGACCGTGAATTTGGGCAAGATC 300
 QY 927 AGCTGTTTAAATTTCTTCAAGAAATGCAAAAGAACTTTGTTGAATTTCTTAAGAAAGA 986
 DB |||||
 301 TCGTGGCTTAAATTTCTTCTGAAGAAATGCTTAAGAACTTTGTTGAATTTCAAAAAAGAA 360
 QY 987 TAGTAAAGAGAT 999
 DB |||||
 361 ACAAAAGTAAGGT 373

RESULT 18
 LOCUS AX951891/c 5718 bp DNA linear PAT 08-JAN-2004
 DEFINITION Sequence 20 from Patent WO03095658.
 ACCESSION AX951891
 VERSION AX951891.1 GI:40782280
 KEYWORDS
 SOURCE Bacillus subtilis
 ORGANISM Bacillus subtilis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 1
 REFERENCE
 AUTHORS bj Rnvad, M.E., j Rgensen, P.L. and Hansen, P.K.
 TITLE Homologous recombination into bacterium for the generation of
 JOURNAL polynucleotide libraries
 PATENT Patent: WO 03095658-A 20 20-NOV-2003;
 Novozymes A/S (DK)

FEATURES
 Source
 1. .5718
 /organism="Bacillus subtilis"
 /mol_type="unassigned DNA"
 /strain="MB1510"
 /db_xref="taxon:1423"

ORIGIN
 Query Match 18.9%; Score 190.8; DB 6; Length 5718;
 Best Local Similarity 55.0%; Pred. No. 6.5e-20;
 Matches 488; Conservative 0; Mismatches 357; Indels 43; Gaps 4;

QY 115 TCAAAAGATGGAGTTGAAATCAAGCACGAGAGAGGTACTACGAAAGTACCTAAACACCTT 174
 DB |||||
 931 TCTAAGATTCAGATGATCCATGATGAGAGAGAGAAACGACAGTAAGCGGCACACCT 872
 QY 175 AAACGTTGTTGTTCTTGTAGTATTCATTGTTGATGCGTTAGTTCGTTTAGATGTTAA 234
 DB |||||
 871 AAGCGGTGGTGTGCTTGAGCTTTCATTCTTGATGCGCTTCAACAATCTCGGCATTAGC 812
 QY 235 CCGTGGGATACGGATGATACAAAAAATCGTATTTATTAACCATTAAGAGATAA 294
 DB |||||
 811 CCGTGGGATCGCAGATGACAAACAAAGATATGATTAAGAAAGTTGTCGCGACGCTC 752
 QY 295 ATTGAAAAATACACTTCTGTAGGAACACGTAAGCAACTTAACTTAGAAGAAATCAGTAAA 354
 DB |||||
 751 ATTG---ATTACATCTGTAGGACACGACGAGACCCCAATCTTGAGGTCACTAGTTC 695
 QY 355 CTTAAACCAAGATTAATTTATGCTGATTAATAATAGACAAAGATTTTATAAGACTTA 414
 DB |||||
 694 TTGAAGCTGATTTAATCATCGCTGACGCTGAGCGCCATAAAAACATTTATATAACAGCTG 635

QY 415 AATAAATTTGCTCTCTAGATTGAATCTGAAAGTTTTCGATGGAGATTATATGAAAAATATT 474
 DB |||||
 634 AAAAAATTCGCCCGAGGATTGAATTTAAAAAGCCGCTGAAGCGACATATGACGAAACGATC 575
 QY 475 GATGCTTTTAAACAAATTTCAAAAGCTTTAGTAAAGAAAGAGGTTAAAAAACGCTTA 534
 DB |||||
 574 GACAGCTTTACGACCATTTGCTTAAAGCATTAATTAAGAGATGAAGGAAAGAAAGCTT 515
 QY 535 GAAGAACACGATTAAGAAATTTGAAGAAATATAAAAAGAAATTAATATGATATAAAATCAA 594
 DB |||||
 514 GCCGAGCACAAAAAGTCAATCAACGATCTAAAGGCCGAATCTCCGAAAGATGAACACCGC 455
 QY 595 AAGTATTTGCTCGAGTAGCTGCTTAATCAGGTTTGTCTCATCCAGCAACTCTTTAT 654
 DB |||||
 454 AACATCGTTCTCGCGTTTGAAGAGCGGATTCCTTCCAGCTTTCATACATCATCATCTAT 395
 QY 655 GTTGTCAATTCCTAAAGTCAACTAGGTTTAAAGAGCAATTAAGTGATGATGTTACTAAA 714
 DB |||||
 394 GACGAGAAATCTTTAAATGCTAGGCTTTACACACGCTGTGAAGTCAGAT----- 344
 QY 715 GGTTTAAGTAAGTATCTTTAAAGGACCTTACTTTACAATGAACACCTGAAACTTTTCTCAA 774
 DB |||||
 343 -----AACGCTTATCAAGAGGTGAGCTTTGACCAATTTGAGCAAA 305
 QY 775 GTGAATCTGAGCGTATGTTCAATATGACAAACAAAGCAAGTTCTTAACGAACTTCACTA 834
 DB |||||
 304 ATCGATCTGATATTTGTTTCTCATCTCGACCAACGAGGCAAAACCAATTTG-----TA 254
 QY 835 AAAGAACTAGAAAGAGATCTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGTTT 894
 DB |||||
 253 GATGAGTGGAAACGAAACCCGCTCTGGAANAATCTCAAGCGGTGNAANAATGGCAAGTC 194
 QY 895 GATATTTTAGACCGTGACTTATGGCAAGATCACTGCTGTTTAAATTTCTTCAAGAAATG 954
 DB |||||
 193 TATGATCGGACCGTGACACTTGGACAGATTCAGAGGCATCAAGCTAGTGAACACAGC 134
 QY 955 GCAAAAGAACTTTGTTGAATTA-TCTAAGAAAGATAGTAAAAAAGATAA 1001
 DB |||||
 133 GCCAAAGATGTGCTTTAAAAAAGTGATATAATAATAGGACAAATGAGAA 86

RESULT 19
 LOCUS D86417 37900 bp DNA linear BCT 18-SEP-1997
 DEFINITION Bacillus subtilis 35.7 kb genomic DNA, 70-73 degree region,
 complete cds.

ACCESSION D86417.1 GI:2443217
 VERSION
 KEYWORDS YfmN; YfmL; YfmK; YfmH; YfmG; YfmA; YfLT; Pel; YfIS; YfLR; YfIQ;
 YfIP; YfLO; YfIN; YfIM; YfIK; YfIL; YfID; YfIC; YfIB; YfIA; YfIM;
 YfM; YfMI; YfMP; YfME; YfMD; YfMB; YfML; YfIJ; YfII; YfIH;
 YfIG; YfIE; YfET; YfES; YfER; YfKO; Trep.

SOURCE
 ORGANISM Bacillus subtilis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 1 (sites)
 REFERENCE
 AUTHORS Yamamoto, H., Uchiyama, S., Nugroho, F.A. and Sekiguchi, J.
 TITLE Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree
 region of the Bacillus subtilis genome reveal genes for a new
 two-component system, three spore germination proteins, an iron
 uptake system and a general stress response protein
 Gene 194 (2), 191-199 (1997)

JOURNAL MEDLINE 97417488
 PUBMED 9272861
 REFERENCE 2 (bases 1 to 37900)
 AUTHORS Sekiguchi, J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-1996) Junichi Sekiguchi, Textile Science and
 Technology, Shinsu University, Department of Applied Biology,
 3-15-1 Tokida, Ueda-shi, Nagano 386, Japan
 (E-mail: jseki@u-gipctc.shinsu-u.ac.jp, Tel:0268-21-5344,
 Fax:0268-21-5331)
 Location/Qualifiers

FEATURES

source	1. .37900 /organism="Bacillus subtilis" /mol_type="genomic DNA" /strain="AC327" /db_xref="taxon:1423" complement (56. .1468) /codon_start=1 /transl_table=1 /product="Trep" /protein_id="BAA22289.1" /db_xref="GI:2443218" /translation="MGEINKSARQIVEAVGGAENIAAATHCVTRLRPAIDESKVDOE MLDQIVKGSFSTNGQOVVIGQTVNKVYAEUVKGTGIGESTKDEVKKASEKNMP LQRAVKTADIFILPAIVTAGLLMGINNLTAREGIFFTSKSVQYVQWADLANMI NLIAATFTFALIGSAVRRFGGNPLLGIVLGMVHPDLLNAWGYAAEQSGEIP VMNIIFGLSEKQVQGVQVLPITLASSYMLAKI EVELTKRTPGEGIQLLVVAPITLLLTGF ASFIIGITPAIGNVLTSGLISVFGSPAALGGLLYGGFYGSALVTGMHHTFLAVDLQ LIGSKLQGTFLWMLALNIAQGSAAAMWFI VKDEKOKGLSLTSGISAYIGITEPAI FGVNLRYRPPETIAMVSSGLAGMYSISQGLVASSVGVGVPGIFSIHMSQYWGAFIGH AIVLIVPAGTAYARFKH" complement (1466. .1483) 1643. .1663 1661. .3202 /codon_start=1 /transl_table=1 /product="YfkQ" /protein_id="BAA22290.1" /db_xref="GI:2443219" /translation="MNRDKALQIKLEDNQDRPISPMLKENLENLTLLTEGCSDIV PROPFGNLCGFVIIRGIVKSEHIQDHARPPLMLHLTDQIDHEEALQNTLSISSV ALETSMKVAASILEGNALVADGHSKGLINIKGGORSIEEPTTESTIRGSREGFT ESLRVNTALVRPRVTFQLKMSIFKIKTKTDVVLAIDGLADPKVIDKAKRIKKI KIDAVLESGYIEEFIEDTYPFPOLQYTERPDTVAALGLGRFAIFTDPTFPVLTGP ITFMQMQASDEYERIMSLRWMFLRYMFLVALYLPAIYAVITVHQDLMPLNMP SVASAREPIPAITEALIMISFEALREAGVRPKTIGQVTSILGALVITGAOVAEAG IVSAPMVIIVSLTGIASTPIRENLAISIRMLRPLPLASIFGIFGIMLGTIILVLH LCKLQSGFIVLSGISPFKRDVKDIFVRAPWMTWTRPGIYSRNGQKGAKREDPKD BENNI" 3166. .3185 3183. .4337 /codon_start=1 /transl_table=1 /product="YfkS" /protein_id="BAA22291.1" /db_xref="GI:2443220" /translation="MKKTIYKCVLPLLICILLTGCWDRTEINDIAFVVSSAIDKKQD YRVAMQIPVLQGGTGGGTAGSKTWYVDSAGTTIREANNKLOTSLSRITNTSH RRTVIGDMARDGAVPVDILTNRPNQRLTALILVSRGEARDILNTDVLQEQPAEM IRELAIITSPVFLSRPMSDLVEIGSDPYAPVISASKTQKGGKSKNLKIDGLAIFK KDRLMDIFKDEHTAALMLLNQARQPEFIVL PNMQGASIQLKSNASFRHAENKGK LSMTIEIRAKGILIMENQSTYETRENDQYIIOKALNRTIKKDVISTVHRLQKLKADPA GFQDRTIRSTATTKNLLKKEWEVYKDMEVHVPIVTIEQGGVLYKTTISH" 4336. .4355 4353. .4553 /codon_start=1 /transl_table=1 /product="YfkS" /protein_id="BAA22292.1" /db_xref="GI:2443221" /translation="MISVIVOTLIVCIAIYAVEMKNFRSANNLTWAFSLIAGSAFL WYWRVNPILLRPLGHLFKYIP" 4563. .4579 4577. .5653 /codon_start=1 /transl_table=1 /product="YfkT" /protein_id="BAA22293.1" /db_xref="GI:2443222" /translation="MDKTSYQGLFPFGALYTLAVGLKHAIPILMIESAKONAWHSYILG VIVITPALWMLHRMLMKHQDNIELSDSPAGRIIILFLSYFLLIHNAIDRFFI NLINILFLPRTPMVLGGVITFVAICIAREGKETITRMAQIFLPFGILVILFLPFLTA TQIEQLNITPVPEGLIPYLGQGYAFGMGELIILPLFSNRSVPKVTIFAILGAL LLAVMLFSSISVFGPNLSTTFDPDAWYIRQIRITDFLDRSDLIITAAFWIPVMVXIA
CDS	GSLLVYVYVGLSFLHSHKIDPKAMVTPTQMFSVVGCFWFFELNTNQLIDFNRIKPIINVVI SLLPLLLIYLIIKSKALFGAKAH" complement (5650. .7065) /codon_start=1 /transl_table=1 /product="YfIA" /protein_id="BAA22294.1" /db_xref="GI:2443223" /translation="MERLLVWIEHISDWLWGPPLIILLTGTGLYFTILLKGPQPRYPL YIFKQTIGSVKSKPKPGEVTPVLPQALTSALSSSTTIGAAINVGPAAIMFGGPGAVFWMM LIALFAIAIKFSESVLAVHYREKNEQGEYVGGPMYITKGLRMKMLGVPFVVALIVEL IPSMVQNGSVSVSLAETFSFNKIYAGIGIAFLGIVIGVIGKIKGVTVWGPARGLYSN AYACAGLLIIVLMNLSSVPFSLVFSNAFTSSSAVGGFAGAALEATVWGPARGLYSN EAGGTAPIAHAAANTDHPVRCQFMSVIGIVDTLIICTTAFIVLASGVMTGKNASN DPAALTAAACQHTFGSGGIFVSVLFFVVTIMVIVFGVKAQAEFLFGRLAGHVIK FVYLALIIIGAAGAKAIWGLDLALVIVVENVIALLLLSRRKVKALYTFEFTSEQYI LKDIRTKQKQPVYTPKEAKNS" complement (7063. .7077) complement (7149. .7175) complement (7190. .7450) /codon_start=1 /transl_table=1 /product="YfIB" /protein_id="BAA22295.1" /db_xref="GI:2443224" /translation="MGYRILKNAGMAEDEGALKKELMTIDHLEKCYDEKEREQLIRK KSEKQMLLDKLVDKKGMFSPASAFYKKNKYVDRLGRNRPSSS" complement (7432. .7584) /codon_start=1 /transl_table=1 /product="YfIC" /protein_id="BAA22296.1" /db_xref="GI:2443225" /translation="MMDFSHIVSDKI KRAIKQGDFOQLPGMGKPLPMTRHTCRNR AWAIVF" complement (7448. .7466) complement (7582. .7596) complement (7600. .7737) /codon_start=1 /transl_table=1 /product="YfID" /protein_id="BAA22297.1" /db_xref="GI:2443226" /translation="MREKVAKNVSESTFRFDITCKTRYLSRNGKIKWYIENCMIKYK V" complement (7735. .7752) 7824. .7840 7838. .9787 /codon_start=1 /transl_table=1 /product="YfIE" /protein_id="BAA22298.1" /db_xref="GI:2443227" /translation="MKTFIKERGLAFFLAVLLWIKTYVGYVNLNIGDINTQIKIL LFVNPSSLPFLGFLKKLQQTALIVHFLMSFLLYANIVYRPFNDPFIPIVI MQAKTNGQLGDSAPSLMRPTDAEYFDTIILILA KVNKPAETSKKSPRIIPASS ILVFLINLAQDSRPELLTRFDNRNLVFKVLGYNFTIYDAVQNIKNSORALADSS DVTENVYKANYDVNNVYFGAEGKNVIYVSLQSFIIDYKIDGKEVTPPLNKL AHDNETYFNFHTQGGKTSDAEFMWENSLYFLAQGSFVNKAQNTLQSPAILKS KNYTSATPHGNTQTFWNRNMYKAEGIDKFDSAYYDMNEENTKNYGMKDPFKRESM PLLESLPQFPVTKFITLSNHPFGQDEGDTDPAGDFDSVVDNVFQSAHYLDQSI EQ FPNDLKDGLYDKSIIIVMYGDHYGISENHNKMAKVLGKDEITDYDNAQLORVLPFIH AAGVGEKVKHTAGDVVDVFTILHLGLVDTKDYMLMSGDILSKHEREVIIPRNGDFIS PKYTKISGYDVTGTGKELDESEVDKSEDSLVKKLEMSDKIINGDILLURLEPVGFKK VNPSDYDTKHDEDSSETSKDNEDK" complement (9789. .9820) complement (9820. .11178) /codon_start=1 /transl_table=1 /product="YfIF" /protein_id="BAA22299.1" /db_xref="GI:2443228" /translation="MLSFLQKLGKSFMLPIAVLPVAVGIILALGREDVFENIPFVYQAGT
CDS	terminator

Query Match	18.9%;	Score 190.8;	DB 1;	Length 37900;
Best Local Similarity	55.0%;	Pred. No. 4e-20;		
Matches 488;	Conservative 0;	Mismatches 357;	Indels 43;	Gaps 4;
Qy	115	TCAAAAGATGGAGTTGGAAATCAAGACACGAAGAAGGTACTACGAAAGTAGTACCTAAACACCCCT	174	
Db	25213	TCTAAGGATTCAGAGTGATCATGACGAAGAAGGAAAAAACGACAGTAAGCGGCACACCT	25272	
Qy	175	AAACGTGTGTTGTTCTTCGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTTAAA	234	
Db	25273	AGCGGCTGTGTTGCTTCAGCTTTCATCTCTGGATGCGTTCACAATCTCGGCATTACG	25332	
Qy	235	CTGTTGGGATACGGGATGATTAACAAAAAATTCGTAATTATTAACCAATTAAGAGATAAA	294	
Db	25333	CCGGTGGGCATCGCAGATGACAAACAAAAAAGATATGATTA AAAAGCTTTGTCGGCAGCTCC	25392	
Qy	295	ATTGAAAATATACACTCTCTGAGAACACGCTAGCAACCTTACTTAGAAGAAATCAGTAAA	354	
Db	25393	ATTG---ATTACACATCTCTGAGCACACGACGGAACCCAACTCTTGAGGTATCATGTTCC	25449	
Qy	355	CTTTAAACCCAGATTTAATTAATTCGTGTAATAATAGACACAAAGGTATTTATAAAGACTTA	414	
Db	25450	TTGAGCCGTGATTTTAATCATCGCTGACGCTGAGCGCCATATAAAAATTTATAACAGCTG	25509	
Qy	415	AATAAAATTTGCTCCTACGATTGAACTGAAAAGTTTTCGATGGAGATTTAATAATGAAAATATT	474	
Db	25510	AAAAAATCGCCCGACGATTGAAATTTAAAAAGCGTGAAGCGACATATGACGAAACGATC	25569	
Qy	475	GATGCTTTTAAAAACAATTTCAAAGCTTTTAGGTAAGAAGAAGAGAGTTAAAAACGCTTA	534	
Db	25570	GACAGCTTTACGACCATTGCTAAAGCATTTAAATTAAGAAGATGAAGGAAAAAGAAAGCTT	25629	
Qy	535	GAGAACAACGATAAGAAAAATTGAGAAATATAAAAAGAAATTAATCTGATAAATAATCAA	594	
Db	25630	GCCGAGCACAAAAGATCATCAACGATCTAAAAGCGGAACCTTCCGAAAGATGAAAACCGC	25689	
Qy	595	AAGGTATTGCTCGAGTAGCTGCTAAATCAGGTTTTGCTTTGCTTCATCCAAAGCAACTCTTAT	654	
Db	25690	AACATCGTTCTCGCGGTTCGAGAGCGGATTCCTTCACGCTTCATCATCATCTCTAT	25749	
Qy	655	GTTGGTCAATTCCTAAGTCAACTAGGTTTTTAAAGAAAGCAATTAAGTGAATGATTTACTAAA	714	
Db	25750	GACGGAGAAATCTTTAAAATGCTAGGCTTTACACACGCTGTAAGTGCAGAT-----	25800	
Qy	715	GGTTTAAGTAAGTATCTTTAAAGGACCTTACTTACAAATGAAACACTGAAACTTTTATCTCAA	774	
Db	25801	-----AACGCGCTATCAAGAGGTCAGCGCTTGAGCAATTTGAGCAAA 25839		

This entry contains data from release R16.1 of the Subtilist database. Further data on gene annotation and detailed information about changes from previous releases can be found at <http://genolist.pasteur.fr/Subtilist/>.

FEATURES

```

source
1. 197409
/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"
/complement(52. .1608)
/gene="yfmJ"
/locus_tag="BSU07420"
/complement(52. .1608)
/gene="yfmJ"
/locus_tag="BSU07420"
/function="unknown"
/feature="similar to ABC transporter (ATP-binding protein)"
/codon_start=1
/transl_table=11
/protein_id="CAB12571.1"
/db_xref="GI:2633066"
/db_xref="GOA:O34512"
/db_xref="UniProt/TREMBL:O34512"
/translation="MSILSVKDLHGFGDRAIFNNVSRLLKGHVGLIGANGRGKST
FMNITIGLEDEGKWSKNRVGVYLDQHTVLEKGSIRDLVKDAFHYLFAMEEEMN
EYNNQGEADPELEKLEEVGVIQDALTNDYFVDSKVEIARGLGSIDIGLERDV
TDSGGQTKVLLAKLEKEPIILLDEPTNLDQHTVLEKGLQYENAFILISHD
IFPLANSVTLHYVNEQLTRYGVGYHOFMEVYVKKQLEAAVKKQOEVAELKDFV
ARNKARVTRNAMSROKLDKMDIELAAEKPKPEFHPKPARTSGKLIPTKDLVIG
YDPSLRPLNRMERGOKIALYGANGIGTKTLLKSLGEIOPLGSGVERGHIHTGVP
EQVKEVNNTCIERVSEPSYQYERAAAPKCGLTTKHIESRVSLSGEBAKVR
LCKLNSETNLLVLVDPTNHLDAKAEKALKEYSKILLISHEPDPFYMDIATETW
NCESWTTKVL"
/feature="yfmJ"
/locus_tag="BSU07430"
/locus_tag="BSU07430"
/feature="yfmJ"
/locus_tag="BSU07430"
/function="unknown"
/feature="similar to RNA helicase"
/codon_start=1
/transl_table=11
/protein_id="CAB12572.1"
/db_xref="GI:2633067"
/db_xref="GOA:O34750"
/db_xref="UniProt/TREMBL:O34750"
/translation="MTQWPLHNAQSFIOENWNASGFKPTPVQEQAAQLIMDGKV
IAESTGTGKTLAYALPVLRIKPEQKHPQAVILAPSRVLVQIFQVQDMKAGSEL
AASLIGGANVKKQVEKKGPHIIVGTGPRFELIKAKKLGKHEVKTIVLDETQDLV
PEHRTMKQIIKTLDRLQLCFSAATKETEDELRLAQPVLKQVRSKAEAGKV
HQLTCDQDKVKLLQKLSLEGMOALVFRDIGNLSVYAEKLAHYHVLGVLHSEAK
KMERAKIATDEGDFPLLATDIAAGLDIENLPYVIHADIPEDGVYHRSRGTRG
KKEGNVLSTVLEBSKLEKAKKLGVLSELNAVYAGGKLT"
2979. .3425
/gene="yfmK"
/locus_tag="BSU07440"
/locus_tag="BSU07440"
/gene="yfmK"
/locus_tag="BSU07440"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB12573.1"
/db_xref="GI:2633068"
/db_xref="GOA:O34536"
/db_xref="UniProt/TREMBL:O34536"
/translation="NASIDRFQVMSPELKILEHWFENEDTRRRMDGMLLDEWVARV
NKDHDVTIMAYDGLQPGMVIIEFGERTYIGLVNPLYRLKGYGKQILQKJMTEDP
TSVREWVACIIEEDNRISLACFAAGTLEDTEPDEGFLTLRN"
complement(3478. .4497)
/gene="yfmJ"

```

```

/locus_tag="BSU07450"
/complement(3478. .4497)
/gene="yfmJ"
/locus_tag="BSU07450"
/function="unknown"
/feature="similar to quinone oxidoreductase"
/codon_start=1
/transl_table=11
/protein_id="CAB12574.1"
/db_xref="GI:2633069"
/db_xref="GOA:O34812"
/db_xref="UniProt/TREMBL:O34812"
/translation="MTASQOQIOLARRPQGVPHEDFRFETIPVPEKQGVLVKTLV
VSDPYMRGMDTKSVVERPALDKALSGGVIAEVVSDGNHLLKGGDIVIGNLSQREFS
AVSSALRKIDTSLAPASAYLIGLMTGLTAYFGLLDIGRPEKETVVSAGAAVGS
TVGQIAKIGARVVVGIAGSDEKIDYKQELQFDEAINYKTADDIQLKALQNAQPGVDV
YFNVGGPISDAVMNLLNEFARIPVCGAIISSYNAESADMDPRVSKLTKTKSLMQV
FIVSDYDRFSEGAQLAEWLKAGLHYEETITEGPEINIPDAFLGPKGENKQGLIK
VSDPS"
/complement(4979. .6199)
/gene="yfmI"
/locus_tag="BSU07460"
/complement(4979. .6199)
/gene="yfmI"
/locus_tag="BSU07460"
/function="unknown"
/feature="similar to macrolide-efflux transporter"
/codon_start=1
/transl_table=11
/protein_id="CAB12575.1"
/db_xref="GI:2633070"
/db_xref="UniProt/TREMBL:O34440"
/translation="MSKGYSSILPSQTTNLGSLYTMVVISFLPKMNSTTIASMTV
LIISIFRFGSAILPLTIKRLKPLTIIISOLIQMLLCLIFYALLRTYSNQTLILVF
IIISCFPFNGWFSPLKSAIIGETISPRDVKANGLLSTVDQTFQFVWGLGGLIAP
LGEGYTIILATIFLLTSLLSLFCPLPHGHANSVIREKSNPKSIVSGWYLPFSQKLR
TIIIMDLIESWAGMIWIGSVSLAFVNEVLHKGESWMGFINGAYLGSIMIGGFIYKLS
ERFONKILNMLIGAVSGSLTIYGFISNSYLALILVLPMPAYILRLDTQETLION
ITTEQTRINIMARSASSLVQTFPMSILAIGAISDFLGLVRLVYVSAGILLVLSAIVGS
QLQPKKKVKNKHSF"
6397. .6468
/gene="yfmH"
/locus_tag="BSU07470"
/locus_tag="BSU07470"
/gene="yfmH"
/locus_tag="BSU07470"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB12576.1"
/db_xref="GI:2633071"
/db_xref="UniProt/TREMBL:O34890"
/translation="MRKLKRNKRSIVSPVLYTNEKW"
6535. .7998
/gene="yfmG"
/locus_tag="BSU07480"
/locus_tag="BSU07480"
/gene="yfmG"
/locus_tag="BSU07480"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB12577.1"
/db_xref="GI:2633072"
/db_xref="GOA:O34722"
/db_xref="UniProt/TREMBL:O34722"
/translation="MTLIMODKSIYNKDIIEILIEVEBELITSVPENKNDKPKSLIDNR
IMTLKIQHPVAENVEGMCIAVPNTYIRVRLIDFVEKMWVPEASVETIINLQDQDP
DLVSFKAMVCANHKIESVAYLSIIDDVRESISYKPKVGLGAQVLSLTLIGV
EKHEBPLVKNNYFQNGLLPNNFDEEKIPOSLLIEFEETEDGKILLIPGFFPEGLN
ENETPDKTFNMKQAVPRQKWLPPFFIDKYPVTNKDYDITFEFIEENHGHIFCHNEPQ
NKQHRNTYMDRDLNHPVTGIDYDAFAYARKGELPTEFQMEKAAARGEKNVWP
WGDKFDPKQVYAGSLYNEPITSLSKSWRENLLKAHADKELNHLTSDIFEQNGESPYGV

```



```
Db 11899 TCTATGTAAGCTGTCGAAATTTGCGCGACAAATCGCGCTCAAAACCTTGBATGCCGATT 11958
Qy 461 ATAATGAAAAATATTGATGCTTTTAAAAACAATTTCAAAAGCTTTAGTTAAAGAAAGAAAG 520
Db 11959 ATCAGGACACGCTTGTATGATCTCTTTCGAATTCGAAAGCGCTCGGCAAGAAAGCGCAA 12018
Qy 521 GTAAAAACGCTTAGAAGACACGATAGAAAATTTGAAGATATATAAAGAAAGAAATAACTA 580
Db 12019 TGGAGAAAAAATTTGGCTGAACATAAAACAAAGCTGGATGAATTTGAAACAGAAATTCGGCA 12078
Qy 581 TGGATAAAAAATCAAAAGGTATTGCTCGAGTAGCTTGAATCAAGCTTGTGCTGCTATC 640
Db 12079 GCCGGAACACAGCATTTCTTCTGCTCGGGAACAACAATGAAGAATCACCGTGGCGATG 12138
Qy 641 CAAGCAACTCTTATGTTGGTCAATTTCTTAAGTCAACTAGGTTTTTAAAGAAAGCAATTAAGTG 700
Db 12139 AA---AAGCTTTTACAGCTCTCAGCTTCTGACGAAAATCGGCTATACATACGCTGTCGGAG 12195
Qy 701 ATGATGTTTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTG 760
Db 12196 ACAGCG---GCAAAAGCGATGCCGAAAACGGTGAAATCCGTCAATATTAATGACGCTCG 12252
Qy 761 AAATTTTATCTCAAGTGAATCTGAGCGTATGTTTCATAATGAC---AAACAAAGCAAGTT 817
Db 12253 AACAGCTGTTGAGAAAGATCCGACGTTTATGTTCTGATGACGGAGAAAAGATTAAG 12312
Qy 818 CTAACGAACCTTCACTAAAGAACTAGAAAAGATCTGTATGGAAGAAATTAACGCTG 877
Db 12313 TCGATGAAGACGGGAAAAGACCGATCGAAAAGATCTCTTTGGAAGAAAGCTCAGCGCAG 12372
Qy 878 TGAATAATCAAGCTGTTGATATTTTAGACCGTGACTTATGGCAAGATCACTGTGTTAA 937
Db 12373 TCAAAAACGGCAAGGTTTATGAAGCCGACAGATTCGCTGGTCTCGCGACGACATG 12432
Qy 938 TTTCTTCAAGAAAGATGCGAAAGAACTGTTGTAATTTATCTAAGAAAGATAGTAAAAAAG 997
Db 12433 ACGGAGCAGATGAGCTGATGATGAATCGATCAGCATCTTCTGCTGAGGATCAABAAG 12492
Qy 998 ATAATAAGTA 1007
Db 12493 AATCTTAGGA 12502
```

RESULT 22

CP000002_11

WPCOMMENT

Sequence split into 43 fragments LOCUS CP000002 Accession CP000002

Fragment Name	Begin	End
CP000002_00	1	110000
CP000002_01	100001	210000
CP000002_02	200001	310000
CP000002_03	300001	410000
CP000002_04	400001	510000
CP000002_05	500001	610000
CP000002_06	600001	710000
CP000002_07	700001	810000
CP000002_08	800001	910000
CP000002_09	900001	1010000
CP000002_10	1000001	1110000
CP000002_11	1100001	1210000
CP000002_12	1200001	1310000
CP000002_13	1300001	1410000
CP000002_14	1400001	1510000
CP000002_15	1500001	1610000
CP000002_16	1600001	1710000
CP000002_17	1700001	1810000
CP000002_18	1800001	1910000
CP000002_19	1900001	2010000
CP000002_20	2000001	2110000
CP000002_21	2100001	2210000
CP000002_22	2200001	2310000
CP000002_23	2300001	2410000
CP000002_24	2400001	2510000

```
CP000002_25 2500001 2610000
CP000002_26 2600001 2710000
CP000002_27 2700001 2810000
CP000002_28 2800001 2910000
CP000002_29 2900001 3010000
CP000002_30 3000001 3110000
CP000002_31 3100001 3210000
CP000002_32 3200001 3310000
CP000002_33 3300001 3410000
CP000002_34 3400001 3510000
CP000002_35 3500001 3610000
CP000002_36 3600001 3710000
CP000002_37 3700001 3810000
CP000002_38 3800001 3910000
CP000002_39 3900001 4010000
CP000002_40 4000001 4110000
CP000002_41 4100001 4210000
CP000002_42 4200001 422336
Continuation (12 of 43) of CP000002 from base 1100001 (CP000002 Bacillus licheniformis D

Query Match 16.7%; Score 168.4; DB 1; Length 110000;
Best Local Similarity 50.5%; Pred. No. 7.7e-17;
Matches 490; Conservative 0; Mismatches 471; Indels 9; Gaps 3;

Qy 41 TTGGCTTATTTGTTTAAATTCGAACTGCAAGTGTGGAATAATAGTTCAAGTAACT 100
Db 10685 TTGCTATTTTATTTTATTTGATGATATACAGTGATGACGGCTGCTCTGGAGGAAGCGAGA 10744
Qy 101 CAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAGGTACTACGAAAG 160
Db 10745 GCACGAAAACCTTCGGTCAGGATTCGGTGACAGTTTACGATGATCTTGGAGAAACGAAAG 10804
Qy 161 TACCTAAACACCCCTAAACGCTGTTGTTGTTTGTAGTATTCATTTGTTGATCGGTTAGTTG 220
Db 10805 TCGCGAATAAACCCGAAAAGTTGTTGTTTCTTGAATTTGGGATTTTATGATCGCTCTTTG 10864
Qy 221 CTTTAGATGTTTAAACCTGTTGGGATAGCGGATGATAACAAAATAATCGTATTTATTAAC 280
Db 10865 ATGCGGGTATTAAAGCTGTCGGAATTCGGAGACGCGGAAACCTTAAGTTTATTACGAGA 10924
Qy 281 CATTAAGAGATAAAATTTGGAATAATACACTTCTGTAGGAACACGTAAGCAACCTTAACCTAG 340
Db 10925 AGGTCAGGGGAAAATCAAGGGGTATACCTTCAGTCGGTTTCGGCGCCCGACCAAGCTTTG 10984
Qy 341 AAGAAATCAGTAAACTTAAACAGATTTAATTTGTTGATGATAATTAATAGACACAAAGTA 400
Db 10985 AAAAAATTCCTTTTAAAGCCGATTTAATTTGCGGATTCGAGCAGGCACACGCGCG 11044
Qy 401 TTTATAAGACTTAAATAAATTCGCTCTACGATTTGAACTGAAAAGTTTCGATGAGAGATT 460
Db 11045 TCTATGATAAGCTGTCGAAAATTCGGCGCAATTCGGCTCAAAAACCTTGAATGCCGATT 11104
Qy 461 ATAATGAAAAATTTGATGCTTTTAAACAATTTTCAAAAGCTTTAGGTTAAAGAAAGAAAG 520
Db 11105 ATCAGGACACGCTTGATGATCTCTTACGATTCGGAAGCCGTCGGCAAGAAAGCGCAA 11164
Qy 521 GTAAAAACGCTTAGAAGAACACGATAGAAAATTTGAAGAATATAAAGAAAGAAATACTA 580
Db 11165 TGGAGAAAAAATTCGCTGAAACATAAACAAGCTGGATTTGAAACAGAAAATTCGGCA 11224
Qy 581 TGGATAAAATCAAAAGGTATTGCTGCGAGTAGCTGCTTAAATCAGTTTGTGCTCATC 640
Db 11225 GCCGGAACACAGCATTTCTTCTCGGGAACAACAATGAAGAATACCCGTGGCGATG 11284
Qy 641 CAAGCAACTCTTATGTTGGTCAATTTCTTAAGTCAACTAGGTTTTTAAAGAAAGCAATTAAGTG 700
Db 11285 AA---AAGCTTTTTCAGGCTCAGCTTCTGACGAAAATCGGCTATACATACGCTGTCGAG 11341
Qy 701 ATGATGTTTAAAGGTTTAAGTAAGTATCTTAAGGACCTTACTTACAAATGAACACTG 760
Db 11342 ACAGCG---GCAAAAGCGGATCCGAAAACGGTGAATCCGTCAATTTAATAATGACGCTCG 11398
Qy 761 AAACCTTTATCTCAAGTGAATCCTGAGCGTATGTTTCATAATGAC---AAACAAAGCAAGTT 817
```

Db 11399 AACAGCTGCTTGAGAAAGATCCGACGTTATTTGCTCTGATCGCGGAGAAAAGATAAAG 11458
QY 818 CTAACGAACCTTCACTAAAGAACTAGAAAAGATCCCTGTATGGAAGAAATTAACACGCTG 877
Db 11459 TCGATGAAGACGGGAAAAGACCGATCGAAAAGATCCCTCTTTGGAAAAGCTCAGCGAG 11518
QY 878 TGAATAATCAACGCTGTGATATTTAGACCGTGCACTTATGGCAAGATCACTGTTTAA 937
Db 11519 TCAAAAACGGCAAGTTTATGAAGCCGACAGATTCGCTGCTCCGACGCGACATG 11578
QY 938 TTTCTTCAAGAAATGGCAAAAGAACTGTTGAATATCTAAGAAAAGATAGTAAAAAAG 997
Db 11579 ACGGACGACATGAGCTGATGATGAATGCATCAGCATCTTCTGCTGAGGATCAAAAGA 11638
QY 998 ATAATAAGTA 1007
Db 11639 AATCTTAGGA 11648

RESULT 23

BSY14084

LOCUS

DEFINITION Bacillus subtilis chromosomal DNA, region 78-80 degrees: aprE to comK.

ACCESSION Y14084

VERSION Y14084.1 GI:2226250

KEYWORDS comK gene; yhfO gene; yhfP gene; yhfQ gene; yhfR gene; yhfS gene; yhfT gene; yhfU gene; yhfV gene; yhfW gene; yhfX gene; yhfY gene; yhfZ gene.

SOURCE Bacillus subtilis

ORGANISM Bacillus subtilis

REFERENCE 1 (bases 1 to 104)

AUTHORS Stahl, M.L. and Ferrari, E.

TITLE Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro-derived deletion mutation

JOURNAL J. Bacteriol. 158 (2), 411-418 (1984)

MEDLINE 84212198

PUBMED 6427178

REFERENCE 2 (bases 1 to 441)

AUTHORS Park, S.S., Wong, S.L., Wang, L.F. and Doi, R.H.

TITLE Bacillus subtilis subtilisin gene (aprE) is expressed from a sigma A (sigma 43) promoter in vitro and in vivo

JOURNAL J. Bacteriol. 171 (5), 2657-2665 (1989)

MEDLINE 89213955

PUBMED 2496113

REFERENCE 3 (bases 10610 to 12556)

AUTHORS van Sinderen, D., ten Berge, A., Hayema, B.J., Hamoen, L. and Venema, G.

TITLE Molecular cloning and sequence of comK, a gene required for genetic competence in Bacillus subtilis

JOURNAL Mol. Microbiol. 11 (4), 695-703 (1994)

MEDLINE 94254725

PUBMED 8196543

REFERENCE 4 (bases 1 to 12556)

AUTHORS Noback, M.A., Terpstra, P., Holsappel, S., Venema, G. and Bron, S.

TITLE Direct Submission

JOURNAL Submitted (26-JUN-1997) Noback M. A., University of Groningen,

MEDLINE Dept. of Genetics, Kerklaan 30, Haren, 9751 NN The Netherlands

PUBMED Location/Qualifiers

1. .12556

/organism="Bacillus subtilis"

/mol_type="genomic DNA"

/strain="168"

/db_xref="taxon:1423"

/map="78 to 80 degrees, aprE to comK"

383. .852

/gene="yhfO"

383. .396

/gene="yhfP"

403. .852

/gene="yhfQ"

/notes="Similarity to several hypothetical proteins, from Haemophilus influenzae (Swiss Prot P44036), Saccharomyces

gene

RBS

CDS

gene

RBS

CDS

terminator

gene

CDS

FEATURES

source

gene

RBS

CDS

cerevisiae (Swiss Prot P39979), a
/citation=[2]
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74538.1"
/db_xref="GI:2226251"
/db_xref="GOA:O07614"
/db_xref="UniProt/TREMBL:O07614"
/translation="MNIRQAKTSDAIAPIFNQYRFFYRQA
HESVILAEENGREFIGTOLYPTFSSVMKRIYILNDLFVVP
DYAONGAKCLTLQTEHNRKARSLEYQNGVEEDTGFVHYCLNV
907. .1916
/gene="yhfP"
907. .919
/gene="yhfP"
924. .1916
/gene="yhfP"
/note="high similarity to yhdH from E.coli (Swiss Prot
P26646), and lower to many alcohol dehydrogenases."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74539.1"
/db_xref="GI:2226252"
/db_xref="GOA:O07615"
/db_xref="UniProt/TREMBL:O07615"
/translation="MSTLFOALQAEKADVDVSVHVKTTISTEDLPKDGVLIVKIVYSGIN
YKDLGAKAGNIVREYPLILIGIDAAGTVSSNDPRPAEGDEVATISVELGVRDGL
SEYASVPGDMLVPLPQNLSEKAMVYGTAGTAALSVRHLEQNGLSPEKSGVLTGAT
GGVGGIAVSMINKRGYDVAVSTGNREAADYILQGLASEVISRVDYDGTILKALSQQW
QGAVDPVGGKQALSLKLIQYGGSAVSGLTGGGEVPATVYFFILRGVSLIGIDSVYC
PMDVRAAWVERMSDDLKPDQLLTIVDREVSLEETPGALKOILQNRIGRIVKL"
2043. .3104
/gene="yhfQ"
2043. .2056
/gene="yhfQ"
2058. .3104
/gene="yhfQ"
/note="Similarity to citrate-dependent iron transport
protein fecB from E.coli (PIR S56515) and Synecocystis
sp. (EMBL: D90899)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74540.1"
/db_xref="GI:2226253"
/db_xref="GOA:O07616"
/db_xref="UniProt/TREMBL:O07616"
/translation="MILPCYTLTGLHNRNYQQRLEHMKTLILITVLLSVLTAA
CSSSSGNSQNSKEHVAVTHDLGKTVNPKRVRVVLGFGITDLDLGTTPVGVADDN
KAKOLINKDLKIDGYTSVCTRSQPSMEKIASLKPDLIADTTRHKVYDQKKIAP
TIALNNLDYQDTIDASLTAKAVGKEKMEKKLTAHEEKLSETKOKISANSQSVLL
IGTNDTITARDENFTSRLLTVQGYRIATISGNSDSSGGSVNMKMTLEQLKTD
PDVILMTGKTDDLDADGKRIEKNVLVKLKAQNGHYHVDRAVSLRRSVSDGANA
ILDELQKEMPAKK"
3099. .3141
complement(3136. .3731)
/gene="yhfR"
complement(3136. .3717)
/gene="yhfR"
/note="similarity to a hypothetical protein o215b from
E.coli (Swiss Prot P36942), and to several
phosphoglycerate mutase (pgm) proteins, from Traponema
pallidum (EMBL: U55214) and Schizosaccharomyces pombe
(Swiss Prot P36623)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74541.1"
/db_xref="GI:2226254"
/db_xref="GOA:O07617"
/db_xref="UniProt/TREMBL:O07617"

/translation="WTAIVLVHGETDWNLOQCKGKTDIPLINATGBERQARETGEYVK
DFSWDIIYVSLPKAKRTAEIINEYLHLPFIVEMDDDFKERYDGDAGMPLEBTRKRYPD
NIYNMETLESLDRLMGLKAVNQAYPNKKVLIVAHGAIAHALLTETISGDPBLOST
RLVNACLSNIIEFAEBEKWRIKDYINSHLSGFIK"

complement(3722. .3731)

/gene="yhfa"

complement(3788. .4882)

/gene="yhfs"

complement(3788. .4882)

/note="yhfS"

/note="similarity to acetyl-CoA acetyltransferases (ThiL,
encoding the enzyme for the first step in
poly-beta-hydroxybutyrate synthesis) from Chromatium
vinosum (Swiss Prot P45369), Alcaligenes eutrophus (Swiss
Prot P14611), Thiocystis violacea (Swiss Prot P45363), and
B. subtilis (Swiss Prot P45855)."

/codon_start=1

/transl_table=1

/product="hypothetical protein"

/protein_id="CAA74542.1"

/db_xref="GI:2226255"

/db_xref="UniProt/TREMBL:O07618"

/translation="MNAVIVDAKTIQNGNLLKPFLEDAAPILIRCLKLEDOV
DEVILGNATGRGNALSAQAGLPLSPGWTIDROCGSGLEAVRVACSLIQAGAGT
MYIAGSSSSQSPSERARSPDAIGDPDMGIAEYTAARYSISRNSQDEYALLSHQ
RSRNAHDEGFYREEVVAGLELDEAFILTRPIEALIPRAKPVFDTSSTGVTVAANSQ
IADGAAALLVMEKEKAAIGLKVLRIFGSAVSGHFNFPFAAPVAIRQLLTHDVT
PDDIDLFEINFAFAVKI CVCSQELGIPFSKINVRGALALGHYPYGAASGAAVLTRIFYE
AKRRPDCOVAVAATIGSGGIGLALLFEVLA"

complement(4879. .6318)

/gene="yhfa"

complement(4879. .6318)

/note="yhfT"

/note="similarity to long-chain-acyl-CoA synthetase from
B. subtilis (EMBL: Z75208), bile acid-CoA ligase from
Eubacterium sp. (Swiss Prot P19409), long-chain-fatty
acid:CoA ligase from E. coli (LcfA; Swiss Prot P29212)."

/codon_start=1

/transl_table=1

/product="hypothetical protein"

/protein_id="CAA74543.1"

/db_xref="GI:2226256"

/db_xref="GOA:O07619"

/db_xref="UniProt/TREMBL:O07619"

/translation="MTIHTYSSTAETSPPRVAIQTESQITVDWDLKVSQTNWLR
SQSPNPNRAVAILLNSLAFLQIFAGAAAAGCTAIPDRWSPAECKERLSISNADLVV
TLAPFKNLTSQTPVLLDNCMADISEAADPLTIDPEHPFVYMGSTGSKPKAF
TRSHRWESPTCTETDFSISSDDKVLIPGALMSSHFLYGVSTLFLGTVCLLKFS
PAKXEWLCRSISVLIVPVTMDALRIEGFPDPSPKVLISSGADWPASCKKLAAAM
PHLKLYDFYGTSELSPVTFSSPEDSKRPHSAGRPVHNRIEIRNAGGERCPQIEGK
IFVKSPMRFSGYVNGSTPDWMTVDMDGYVDEEGLFYISGRENGMIVYGGNLNIPFEEI
ERVLLACPEVSAAVVGIPDEYVGVIAVAVILGNANARTLKWCKQKLASYKIPKKW
PADSLPSTSGKIAARSVRKKWLEESVQK"

complement(4893. .4903)

/gene="yhfs"

complement(4893. .4903)

/note="yhfS"

complement(join(6325. .6885,6888. .6897))

/gene="yhfu"

complement(6325. .6885)

/note="similarity to BioY from Bacillus sphaericus (Swiss
Prot P22819)."

/codon_start=1

/transl_table=1

/product="hypothetical protein"

/protein_id="CAA74544.1"

/db_xref="GI:2226257"

/db_xref="UniProt/TREMBL:O07620"

/translation="MLKIIDMMHIAIFTALMAVLGMPPLFLSFTFPVITITLQVLML
AGSILRPSAFSLQVLVLLVAFKAPLLPGGSGRGVFGPSAGFLIAYPLASWLISL
AANRLKVTVLRLEPFTHIVFGIIFLYLLGIPVQAFIMHIDLSQAAPMSLAVVPGDLIK
RAVSAFLAIKIQALSIDTWTPTKGG"

gene complement(6329. .6336)
/gene="yhft"
RBS complement(6329. .6336)

Query Match 16.5%; Score 166.6; DB 1; Length 12556;

Best Local Similarity 51.5%; Pred. No. 2.5e-16;

Matches 460; Conservative 0; Mismatches 424; Indels 9; Gaps 3;

Qy 117 AAAAGATGGAGTTGAAATCAAGACGAGAAAGGTACTACGAAAGTACTCTAAACACCTTAA 176

Db 2222 AGAACATAAAGTGGCGGTAAACATGATTTAGGGAAGACAAATGTGCTGAGCATCCGAA 2281

Qy 177 ACGTGTTGTTCTCTCAGTATTCTTTGTTGATGCGTGTAGTCTGCTTTAGATGTTAAACC 236

Db 2282 GCGGGTGTGTTCTTCTGAGCTAGGTTTATTGATACACTGCTTGTATCTCGGCATTAACGCC 2341

Qy 237 TGTGGGATAGCGGATGATTAACAAA---AAAAATCGTATTATTAAACCATTAAGAGATAA 293

Db 2342 TGTGGGGTTGCCGATGACACAAAGCGAGCAGTATCAACAGGATGTGCTGAAGAA 2401

Qy 294 AATTGGAATAATACACTCTCTGTAGGAACACGTAAGAAACCTTAACCTTAGAAGAAATCAGTAA 353

Db 2402 AATTGACGGCTACACATCTGTGCGCACTCGCTCACAGCCAGCATGGAAGAAATTCGCTTC 2461

Qy 354 ACTTAAACACAGATTTAATTATTTGCTGATAATAATAGACACAAAGGTATTTATAAGACTT 413

Db 2462 ATTAACACCGGATTTAATTATTTGCTGACACACCCCGCATAAAGAGGTGTACGATCAGCT 2521

Qy 414 AAATAAAATCTCTCTACGATTGAACCTGAAAGTTTTCGATCGAGATTATTAATGAATAATAT 473

Db 2522 GAAAAAATAGCGCGGACGATTGCACITTAATAATTTAAATCTGATTATCAGGATACAAAT 2581

Qy 474 TGATGCTTTTAAAAACAATTTCAAAAGCTTTAGGTAAAGAAAGAAAGTAAAAACGCTT 533

Db 2582 TGACGCTTCGCTTACGATTGCAAAAGCAGTCCGCAAGGAGAAAGTAAAGGAGAAAGCT 2641

Qy 534 AGAAGAACACATAGAAATTTGAAGATATAAAAAAGAAATTAATATGGAATAAAAAATCA 593

Db 2642 GACGCGCATGAAGAAAGCTTTAGCGAGACAAAGCAGAAATTCAGCGGAACAGCCAGCTC 2701

Qy 594 AAAGTATTGCTGCGAGTAGCTGCTAAATCAAGTTTGTCTTCATCCAAAGCAACTCTTAA 653

Db 2702 CGTGCTTTTGATCGGAATAACAAATGATACCAATATATGCCAGGATGAAACTTCTTTAC 2761

Qy 654 TGTGTCCTCAATCTCTAGTCAACTAGGTTTAAAGAGCATTAAGTATGATGTTACTAA 713

Db 2762 ATCGAGACTTT---TAACACAGGTCCGCTACCGGATATGCAATTCAGTACGTACGCAATAG 2818

Qy 714 AGGTTTAAAGTATCTTTAAAGGACCTTACTTCAAAATGAACACTGAAACTTTTATCTCA 773

Db 2819 CGATTCAAGCAATGGCGGCGACTCTGTGAATATGAATAATGACACTGGAGCAGCTGCTGAA 2878

Qy 774 AGTGAATCCTGAGCGTATGTTTATAATGACAAAAAAGCAAGTCTTCAACGAACTTCACT 833

Db 2879 AACAGATCCGATGTGATCATCTCTGATGACAGGAAAAACAGATGACCTCGACGCGCG 2938

Qy 834 AAAAGAAC---TAGAAAAAGATCTCTGTATGAAGAAATTAACCGTGTGAAAAATCAACG 890

Db 2939 TAAACGCCGATCGAAAAAGAAATGTCTTTTGAAGAAATCTGAAGGCGAGTGAAGAAACGGCA 2998

Qy 891 TGTTCATATTTAGACCGTGACTTATGGCAAGATCAGCTGTTTAAATTTCTTTCAGAAGA 950

Db 2999 TGTATACACGTGGATCTGCGGTGCTCTCTGCGCGCGAGTGTAGACGGGGCGAATGC 3058

Qy 951 AATGGCAAAAGAACTTGTGTAATTTATCTTAAGAAAGATAGTAAAAAGATAATA 1003

Db 3059 CATTTTGACGAGCTTCAAAAAGAGATGCGCGGCTGCTAAGAAATAAAGAAAAA 3111

RESULT 24

BSUB0006

LOCUS

DEFINITION Bacillus subtilis complete genome (section 6 of 3

BSUB0006 198743 bp DNA linear

to 1209781.
299109 AL0009126
VERSION 299109 2 Gi:32468723

Bacillus subtilis subsp. subtilis str. 168
Bacillus subtilis subsp. subtilis str. 168
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

1 (bases 1 to 198743)
Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S.,
Boriss, R., Brouzier, L., Brans, A., Braun, M., Brignell, S.C.,
Bron, S., Brouillet, S., Bruchi, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Connetton, I.F., Cummings, N.J.,
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E.,
Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A.,
Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Gollightly, E.J.,
Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haeche, J., Harwood, C.R.,
Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hulio, M.F.,
Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaer-Bianchard, M., Klein, C., Kobayashi, Y., Koetter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P.,
Mizuno, S., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,
Portetelle, D., Porwollik, S., Prescott, A.M., Prescan, E., Pujic, P.,
Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadate, Y., Sato, T.,
Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B.,
Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpatra, P., Tognoni, A.,
Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A.,
Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenecker, T.,
Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and
Danchin, A.

The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
Nature 390 (6557), 249-256 (1997)
98044033
9384377

2 (bases 1 to 198743)
Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
Direct Submission
Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
Genetique des Genomes Pasteuriens, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
danchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

On Jul 7, 2003 this sequence version replaced gi:2633260.
This entry contains data from release R6.1 of the Subtilist
database. Further data on gene annotation and detailed information
about changes from previous releases can be found at
<http://genolist.pasteur.fr/Subtilist/>.

Location/Qualifiers
1. 198743
/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"
complement(38..1504)
/genes="lytF"
/locus tag="BSU09370"
complement(38..1504)
/genes="lytF"
/locus tag="BSU09370"
/function="cell wall lytic activity"
/note="alternate gene name: cwle, yndd"
/codon start=1
/transl table=11
/product="gamma-D-glutamate-meso-diaminopimelate

muopeptidase (major autolysin) (CWBP49')"
/protein_id="CAB12776.1"
/db_xref="GI:2633272"
/db_xref="GOA:O07532"
/db_xref="UniProt/TREMBL:O07532"
/translation="MKKKLAAGLTASAIVGTTLVTPAEATATIKVSGDSLWKLAKQTY
NTSVAULTSANHLSTTVLSIGQTLTIPGSKSTSSSTSTTKSGSSSVVTKSGDSL
WLIANEFKMTVOELKGLNGSSDLIRAGOKLVSGTVSSSSSSKKSSNKSNKSVKSSSK
SSSNKSSSSSTGTYKVQVGLSLWKIANKVNSIAELKVLNKLKSDTIYVNOVLTKK
SSGSDTSKONSSKSNSTSAITKITVKSGLSKIANNYNLTVQOIRNINNLKSDVLY
VQVGLKLTGKASSGSS
NNLKTDLVQGLKIVISGKASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
TTPSGDCSGFIYYVLNKTYSVRLTAAGYNTWKVSQPAVGDFVFVFSTYKAGPSHV
GIYLGNEFTINANDSGVVISNMNNSYKQRYLGAKRYF"

gene
/gene="yhdE"
/locus tag="BSU09380"
complement(1657..2097)
/gene="yhdE"
/locus tag="BSU09380"
/function="unknown"
/note="similar to unknown proteins"
/codon start=1
/transl table=11
/protein_id="CAB12777.1"
/db_xref="GI:2633273"
/db_xref="GOA:O07573"
/db_xref="UniProt/Swiss-Prot:O07573"
/translation="MKLTNTDYSRLVIFLAERPGELSNIKQIAETYSISKNHLMKV
VIYRLGQVGYETIRGVLGMDPDINIGEVKRTEDDFNIVECFDANKNLVCI"
SPVCKLHVLEALLAVLDKYTLRDLVKNKEDIMKLLKMKKE"

gene
/gene="yggB"
/locus tag="BSU09390"
complement(2200..3858)
/gene="yggB"
/locus tag="BSU09390"
/function="unknown"
/codon start=1
/transl table=11
/protein_id="CAB12778.1"
/db_xref="GI:2633274"
/db_xref="GOA:P37874"
/db_xref="UniProt/Swiss-Prot:P37874"
/translation="MYEPARQKIVISGQEKIPAHSSVCLLDKWNTHHTTKERYQV
LNADVFSYLSKLPNIIALLVLLIGAKITAEKAYKGLSKTKTDDKLKFKGKPS
RYSSEKTSKVYFIALIIVFIFLFTNLHTTVASPFVMSLSAIAAIPSVLKAGLLI
LLGWAASVLSFLVKIKGMKLTSDKLKRNVLVSEKDIHQAVNTASQIVFYLVLVFL
LPQVLSUKISGISGPFNTMMESVLAFLPKLFAALVILIGWLVARLVDRDIITNFLAS
IGTERFARMGLSIYKDTLSUAVIGTITAYVILIMPVVISALDQDVLVAGISKPAKUS
NTILNLMIMIAIVLVLGAGWGVKVMVSLGHRGFDVSLGRKMGAGTTPAKVLS
LSQVGMIAQIIVILLFTAEALQIVRIHLFELVATGIIAYLPNVLVAVFLLGLGYSF
ELVRVLTASMIKGQEFKSLAPKIATYIIALFAFFMALDQLGVAATVNSAFIIVLSGFA
LAFGLSFLGCKGFASRVLSLTFERKMQNTEIEKRNKNQNPNDM"

3889..5295
/gene="spoVR"
/locus tag="BSU09400"
3889..5295
/gene="spoVR"
/locus tag="BSU09400"
/function="involved in spore cortex synthesis (stage V
sporulation)"
/codon start=1
/transl table=11
/protein_id="CAB12779.1"
/db_xref="GI:2633275"
/db_xref="GOA:P37875"
/db_xref="UniProt/Swiss-Prot:P37875"
/translation="MNAEEKQLQRAIBEITEIARGLDFPMPRYICPABIIYTFG
AYGMPTRFHSWFGQFHKMLHYDLGLSKIYELVINDSPCYAFPLNLSNLIQNLVIA
AHVLAHCDFFKNNCFONTKEDVMVESMAATAERIKHYEOVHGIKEVESFDAIISIOE
HIDPSLVRPQKQWSDDEEIEEAATPYDILNSDEKKKQVKKSKKPPPPPEKDK
ILLTFEHSRELPQWRDITWMEEMFWPQLETIKWNEGWSAYHQRIIRSDUDT

FEATURES
source
gene
CDS

CDS	/gene="PM0119" complement(3197..3739) /gene="PM0119" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAK02203.1" /db_xref="GI:12720332" /translation="MAKRGKNIENDTWDQBEHEIIWVSKSEIKRDAEALKKLGK LVDTAKLEPIEKELEAIEAQRLOKEARRRQYIGIKLRNIDVEPIQALEK LENHQQQOAMLHKLEMLRDLIDNPEAITQLEQHPDADRQHLNLRVAAQKEKAQ KPKYAIRIIFOVLKDFLED" 3834..5189 /gene="pmba" /note="synonym: PM0120" 3834..5189 /gene="pmba" /codon_start=1 /transl_table=11 /product="Pmba" /protein_id="AAK02204.1" /db_xref="GI:12720333" /translation="MQNTEKSTALLKQEQTLREAVRYAVDLAQKAGATAEAVTVKS GLSVTRLOETENVEFTWDGALGIVLGNQGNASTGDLSPQAKNLYVEAKIISAKY ASPDACGLADYDLWAPDAPDLALVHAADIDVEKATALATEQAALSYDKRIVNSEG SAFNSHTGVRYGNTHGMLQSLRSYSLSCSVIAGVEDQLERDYEYTI SRDLAQLES ASWGEHCAKKTLSRLQPKMAQQPVIFLNDVATGLISHLTAISIGSLYRKSSEL LDHGKGLPWFHISEQPHLIGQIASSPFBSEGVTRQCREIIQQGTLOTYLLTSYG KALGRSTAGGAGIHNLVKNPTTGLTTLRLQMGTLGLVTLDMGQQVNLVTGYSRG AAQPVWENGEIQYPAEITTIAGQLKEMLRNIVA VGDDIETHRSNIQTGSLLEHLKVS G" 5424..5963 /gene="hpt" /note="synonym: PM0121" 5424..5963 /gene="hpt" /codon_start=1 /transl_table=11 /product="hpt" /protein_id="AAK02205.1" /db_xref="GI:12720334" /translation="MKHHVGILISEQDVRAIRIOTLGREITQYQOKAVEKIVVVGLL RGSFMDLVRELNLPEIEFMTTSSYSGWTTNHDVKISKLDGDIKGBHVLIED LIDGTYLOKREILKREPASLRILCTILDKPSEVEVEPVWVGFSIPDEFVVGVI DIAQYRNILGYIHGVLEE" 5996..6580 /gene="gmha" /note="synonym: PM0122" 5996..6580 /gene="gmha" /codon_start=1 /transl_table=11 /product="Gmha" /protein_id="AAK02206.1" /db_xref="GI:12720335" /translation="WYLSQKAELEAATVNLKFNENNMKLIOQALLISDSFKQG KVLSCGSGSHCDAWHAEELTGRYRNPFGYPAIASDVSHLSVCVNDGFYEVFS RYVEAVGQGDVLFGLSTSGNSKILNIAKAEKGMKVIAMTGDGKGQOMAGIADVEI RVPHFGYADRIQEIHKIVIHILMLIPEMAKTE" 6756..7487 /gene="artp" /note="synonym: PM0123" 6756..7487 /gene="artp" /codon_start=1 /transl_table=11 /product="Artp" /protein_id="AAK02207.1" /db_xref="GI:12720336" /translation="MAIRKVNLFYFGSQALFDINLSABEGDVTVLLGPSGAGKSTL INTNLLEVPAGEELIANSQDSETKTNPKQICRLRQDVGMVFQQVHLMPLHVLIE NLIEAPKVLKSDQASEALELKRLELAQFADRFPLHLHSGQQQQRVATARALMMK POVLLFDEPTALDPEITAVVDIIKELQQTGITQIVTIVHEVGAKVATKVVYMEQG	gene	HIEMGDASCFTNPKTQFKQYLSH" 7504..8226 /gene="arti" /note="synonym: PM0124" 7504..8226 /gene="arti" /codon_start=1 /transl_table=11 /product="Arti" /protein_id="AAK02208.1" /db_xref="GI:12720337" /translation="MKKLLTALLASTALAAAKEITFAMEPSYPPFELTNEKGEIIG FDVDIANAICEIQATCHFQQAQFDSLITGLRSKHFDAASAIIDITBARAKQVSPSDP YDSSSFISVKGKVDLQAOKKVGONGTYYQYIVAEAKQYQVPVPYASIQNALIDL KNGRIDVIFGDTAVLVDMLAKEQSLAFVGDKVTNKYFGNGFGIAVNKASKDLDDL KGLATIRANGEYIKIYDNMTK" 8233..8904 /gene="artQ" /note="synonym: PM0125" 8233..8904 /gene="artQ" /codon_start=1 /transl_table=11 /product="ArtQ" /protein_id="AAK02209.1" /db_xref="GI:12720338" /translation="MFEDYLSLWYATLMTLGLAVCSILVGLIILSIIPVLETSKWOM IRKTSVFTLLRGLPEILLVLLIYFGSTEVEMLTGEFQFSFGVAFALSIIFAA YASQSLRGATQAIPLQWESGAALGRGYTFVQIIMPQVWRHALPGLSNOMVLKDK TALVSLIGVEDLMRQAEVLNVTHTQPTFYGIAALYLTITLISQFIIRKLEPRTFR ERGV4" 8904..9587 /gene="artM" /note="synonym: PM0126" 8904..9587 /gene="artM" /codon_start=1 /transl_table=11 /product="ArtM" /protein_id="AAK02210.1" /db_xref="GI:12720339" /translation="MFQDYPMVIAQGIPTSLLLTFVSLVAFVALLLTFGLSIGNKL KCGTHYLYLTFTQPLLVQFLLIYAGQGPQEWLVDSALMPLLSNMFCAMLTLAINS VACSTHLYLFGHAKAISQWESCAALGLTRMETLKILIPYALKRALPSYSNEIILVFK GTALASITIMDINGYARQLYGYTEYDALTIVIGIAGMIYLIITATATLLLRKLESKVL FERLEVEKA" complement(9646..9921)	gene	Query Match Best Local Similarity 15.5%; Score 156.6; DB 1; Length 13792; Matches 456; Conservative 0; Mismatches 384; Indels 39; Gaps 3; QY 123 TGGAGTTGAAATCAAGCAGGAAGGTTACTACGAAAGTACCTAAACACCTTAAACGTTGT 182 DB 13549 TGCTGTTACCGTCAAGATCAAAAAGGGGAATTTACCTTAGATAGCGTACCTTAAACGTTGT 13490 QY 183 TGTGTTCTTCAGTATTCATTTGTTGATCGCTAGTTCCTTTAGATGTTTAAACCTGTTGG 242 DB 13489 TGTTCCTTGAAGATATTTCTTGTGGATGCACTTGCACAAAATGTTGTGTCGCCCTGTTGG 13430 QY 243 GATAGCGGATGATAACAAAAAATTCGTATTATTATAACCATTTAAGAGATAAAATTTGAAA 302 DB 13429 TGTCCAGATGACATGATAAACCCTGTTCTGCAAAAGTACGCGATAAGTCAGCC 13370 QY 303 ATACACTTCTGTAGGAACACCGTAGCAACCTTAACTTAGAAGAAATCAGTAACTTAAACC 362 DB 13369 ATGGGAATCAGTGGGACCGCTTCTCAACCGAGTTTAGAAGCGATTTCTGCACTTAAACC 13310 QY 363 AGATTTAATTATTCGTGATAATAATAGACACAAAGGTTATTTATAAGACTTTAAATAAAT 422 DB 13309 AGATTGATCATTCGCCGATGATATCGCCATCTCCGCTCTATGAGAACTCAAAAAT 13250 QY 423 TGCTCTCAGATTGAACCTGAAAAGTTTCGATGGAGATTATAATGAAATATTGATGCTTT 482
-----	---	------	--	------	--

```

13249  CGCGCGCAGTCGCTTTAAATTCGCCCATGAGAACTATCAAGAAACCTTTGAAACCGC 13190
483  TAAACAATTTCAAAGCTTTAGGTAAGAAGAAGAGTAAGAAACGCTTAGAAGACA 542
13189  ACAAAAATCGGTGATTTATTAGGTAAATCAAAAGAAATGCAAGCGCGTATTGCAAAACA 13130
543  CGATAAGAAAATTTGAAGATATAAAGAAAGAAATACTATGGATAAATAATCAAAAGGTATT 602
13129  TAAGCAGGATATTGGGAC-----ATGCCAAAGCGTTACGAAAGGGAAGAAACCGAT 13076
603  GCCTGCAGTAGCTGCTTAATCAGGTTTGCTCATCCAAACCACTCTTATGTTGGTCA 662
13075  TATCGGTGTTTCAAGTGAACCCAAATTTAATTTATATATATAGCAATCTATGCTGGTGG 13016
663  ATTCCTAAGTCAATCAGGTTTAAAGACGATTAAGTATGATGTTACTAAGAGTTTAAG 722
13015  CTTAGTGAAGTCTAGGTTTATCAATGCCAAAGCCCGTGGGATTAACCAACCTAATGC 12956
723  TAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCC 782
12955  TTCGGTGTGTTAGAACAA-----GTGGCGGCGAAGAAAGCC 12920
783  TGAGCGTATGTTTCATATGACAAACAAAGCAAGTTTCTAACGACCTTCTCACTAAAGAACT 842
12919  TGATCTGATGATCTTAATCCATTA-----TCGTGATGAGAGTATTGCAAGAAATG 12869
843  AGAAAAGATCCTGTTATGGAAGAAATTAACGCTGTGAAGAAATCAACGTTGATATTTT 902
12868  GGAAAATGAAGCCTTATGAAAATTTATCCCTGCGGTAAAGAAATGGTCAAGTCTATCTAGC 12809
903  AGACCGTGACTATGGCAGAGTACAGTGTGTTTAAATTTCTTCAGAGAAATGCAAGACA 962
12808  TAATGATATTTATGGCAGAGACAGTGTGTTGATGCGCTGGAAGTAAATGGCTTAAGAA 12749
963  ACTTGTGTAATTTATCAAGAAAGATAGTAAAAAGATAA 1001
12748  AGTCCAGACTTTGTAAACGAATCCGCCAAATTAATGAAA 12710

RESULT 28
AX433651      801 bp      DNA      linear      PAT 28-JUN-2002
LOCUS
DEFINITION   Sequence 2066 from Patent WO0229113.
ACCESSION   AX433651
VERSION      AX433651.1 GI:21658459
KEYWORDS     Bacillus licheniformis
SOURCE       Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ORGANISM
REFERENCE    1
AUTHORS      Berka, R. and Clausen, I. G.
TITLE        Methods for monitoring multiple gene expression
JOURNAL      Patent: WO 0229113-A 2066 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES     Location/Qualifiers
             1..801
             /organism="Bacillus licheniformis"
             /mol_type="unassigned DNA"
             /db_xref="taxon:1402"

ORIGIN
Query Match 14.8%; Score 149.4; DB 6; Length 801;
Best Local Similarity 51.8%; Pred. No. 2.1e-13;
Matches 414; Conservative 0; Mismatches 376; Indels 9; Gaps 3;

QY 176 AACGTGTGTTGTTGATGATTTCAATTCGTATGCGTTAGTTGCTTTAGATGTTAAAC 235
DB 1 AAAAAGTTGTGGTCTTGAATTTGGAATTTATGATGCGCTCTTGTATGCGGGTATTAAAC 60
QY 236 CTGTGGATAGCGGATGATAACAAAAAAATCGTATTATTAACCACTTAAGAGATAAAA 295
DB 61 CTGTGGAAATTCGGACGACGCAACCTAAGTTTATTAACGAGAGGTTCAGGGGAAAA 120
```

```

QY 296 TTGGAATACACTTCTGTAGGAAACAGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAC 355
DB 121 TCAAAGGGTATACCTTCAAGTCGGTTCGCGCCAGCCAAAGCTTTGAAAAAATTTGCTCTT 180
QY 356 TTAACCCAGATTTAATTTCTGCTATATATAGACACAAAGGTATTTTATTAAGACTTAA 415
DB 181 TAAAGCCCGATTTAATTTATTCGGATTCGACGAGCAGCGGCGTCTATGATAAGTGT 240
QY 416 ATAAATTTGCTCTACGATTCGAACTGAAAAAGTTTCATGGAGATTTATATGAATAATTG 475
DB 241 CGAAAAATTCGCGCGCAATTCGCGTCAAAAACTTGAATGCCGATTTATCAGACACGCTTG 300
QY 476 ATGCTTTTAAACCAATTTCAAAGCTTTAGGTAAAGAAAGAGAGGTAAAAAACCCCTTAG 535
DB 301 ATGCACTCTTACGATTCGGAAGCCGTCGCAAGAAAGCGCAATGAGAAAAAATGG 360
QY 536 AAGAACACGATAAGAAATTCGAAGATATAAAAAAGAAATAACTATGATTAAGATCAAA 595
DB 361 CTGAACATAAAACAAAGCTGGATGAATTTGAACACAGAAATTCGGCAGCCGGAACAGACA 420
QY 596 AGGATATTCCTGCGAGTACTGCTAAATCAGGTTTGTCTCATCCAAAGCAACTCTTATG 655
DB 421 TTCCTCTCTCGGGAACAAATGAAGAAATCAACGTCGCGCATGAA---AACTTTTCA 477
QY 656 TTGCTCAATTCCTAAGTCAACTAGGTTTAAAGAAAGCATTAAAGTGATGATGTTACTAAG 715
DB 478 CGTCTCAGCTTCTGACGAAATTCGCTATACATACGGTGTGCGGACAGCG---GCAAG 534
QY 716 GTTTAAGTAAGTATCTTAAAGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAG 775
DB 535 GCGATGCCGAAAAACGGTGAATCCGTCATATTTAAATGACGCTCGAACAGCTGCTTGAGA 594
QY 776 TGAATCCTGAGCGTATGTTTCAATGAC---AAACAAAGCAAGTTCTTAACGAACCTTCA 832
DB 595 AAGATCCCGACGTTATGTCCTGATGACGGGAAAAAAGATAAAGTCGATGAAGACGGA 654
QY 833 TAAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTG 892
DB 655 AAAGACCGATCGAAAAAGATCCTCTTTGAAAAAGCTCAGCGCAGTCAAAAAACGCAAGG 714
QY 893 TTGATATTTTAGACCGTGAATTCATGGGCAAGATCAAGTGGTTTAAATTTCTTCAAGAAA 952
DB 715 TTTATGAAGCCGACAGATTCGCTGCTCGCTCCGACGACGATTTGACGGAGCAGATGAGC 774
QY 953 TGGCAAAAGAACTTGTGTA 971
DB 775 TGATGGATGAATCGATCA 793
```

```

RESULT 29
AE017012
LOCUS
DEFINITION   Bacillus cereus ATCC 14579 section 15 of 18 of the complete genome.
ACCESSION   AE017012 AE016877
VERSION      AE017012.1 GI:29897926
KEYWORDS     Bacillus cereus
SOURCE       Bacillus cereus ATCC 14579
ORGANISM     Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE    1 (bases 1 to 301332)
AUTHORS      Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapral, V., Bhattacharya, A., Reznik, G., Mikhailova, N., Lapidus, A., Chu, L., Mazur, M., Coleman, E., Larsen, N., D'Souza, M., Walunas, T., Grechkin, Y., Pusch, G., Haselkorn, R., Fonstein, M., Ehrlich, D. S. D., Overbeek, R. and Kyrpides, N.
TITLE        Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis
JOURNAL      Nature 423 (6935), 87-91 (2003)
MEDLINE      22608415
PUBMED       1271630
REFERENCE    2 (bases 1 to 301332)
```

AUTHORS	Candelon,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.
TITLE	The number of ribosomal RNA operons in <i>Bacillus cereus</i>
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 301332)
AUTHORS	Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B., Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goitsman,E., Larsen,N., D'Souza,M., Walunas,T., Greckin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France
FEATURES	Location/Qualifiers
source	1..301332
	/organism="Bacillus cereus ATCC 14579"
	/mol_type="genomic DNA"
	/strain="ATCC 14579"
	/db_xref="ATCC:14579"
	/db_xref="taxon:226900"
	complement(114..470)
gene	/locus_tag="Bc4288"
CDS	complement(114..470)
	/locus_tag="Bc4288"
	/codon_start=1
	/transl_table=1
	/product="Membrane-attached cytochrome c550"
	/protein_id="AAP11201.1"
	/db_xref="GI:29897927"
	/translation="MKRNPLIPFALIAALGIIVMFVSFOGLINKSEKLADANGKPAQTASPEDIVQSCSTCHGDQLQAGVGNLQIGGKLSKDEIKELSKGNMNPANVPADQAQKADNLSKKK"
gene	complement(857..1984)
CDS	/locus_tag="Bc4289"
	complement(857..1984)
	/locus_tag="Bc4289"
	/codon_start=1
	/transl_table=1
	/product="RNA polymerase sigma factor rpoD"
	/protein_id="AAP11202.1"
	/db_xref="GI:29897928"
	/translation="NADKPARSKQIETEMTLQVKEQLTELKKGKGVLTVEIAERNNGFEISDDMDVEYVLGQGLDVGNDNEEGPNRQISKTDEEPDLNDLSVPPGVKINDPVRMYLKEIGRVDLLSAREIRLATRIESGDEAKRLAEANLRLVSTAKRVGHRGMLFDLLIQBNMGLIKAVEKFDYRKGFSTYATWIRQALTRAIADQARTIRIPVHMYETINKLIRVORQLDGLRPEPSPEEIGEEMDLAPEKVRTELKTAEPVSLTEPIGEDSHLGFIEDQATSPADHAAYELLKEQLEDVLTLDREENVLRRLRFLDGDGRTITLEEVKGVGVTRIRIQIEAKALKRLRHSRSKELKDFLE"
gene	complement(2044..3840)
CDS	/locus_tag="Bc4290"
	complement(2044..3840)
	/locus_tag="Bc4290"
	/EC_number="2.7.-"
	/codon_start=1
	/transl_table=1
	/product="DNA primase"
	/protein_id="AAP11203.1"
	/db_xref="GI:29897929"
	/translation="MGNRIPEEVBEQIRTSDDIVRIGEYVQLRKQGRNYFGLCPFHGENSPFSVSDSKQIFHCFCGEGNVFSLPMKEGLAFTEAVQKLGERNGIAVAEYTSGGQOEDISDDTVIMQQAHELKYYHLLVNTREGNALS YLLNRGITKMEIKFEI GYASPAWDAATKILQKGLSLTMOAGLLIRSEKDGSHYDFRGRVMFPIYTLQKV TAPSRALGDDTPKYLNSPETPIFKSKLLYNFHOARPPIRKGGQVLFEGYADVLA AVKSGVEAVATGTALEQAKLRNVETVVICYDGDKAGREATKAGQLLKVKCCQ VKVTSPLDKLDDEYVQGYTTAFENLVKSSVFGFKINYLRLGNLODESGKEEYV KSVLKELSLLDQMAQESYLSQSOFYSMETLNLQHLQYRKEQVQKQIKQVAKP SOVVTQPKLTGFERAEREIIVHMQSAEVAVRMESHIEDFTEEHKGLLYELAYYE KGNESVGTFLSLSDEKLNIIITDSTDERINPEYTEVQLGHLETLRRHQEKLEKM EIIFKVKMEKDTPEAAKYVYAYLQNKARK"
gene	4007..4213
CDS	/locus_tag="Bc4291"
	4007..4213
	/locus_tag="Bc4291"
	/codon_start=1

	/transl_table=1
	/product="hypothetical protein"
	/protein_id="AAP11204.1"
	/db_xref="GI:29897930"
	/translation="MLYKIVYICSSVTNVSRLRNNTYTPFYKSKSPHILFDLEICNFFMLHITNSILFSYKDPDIALFP"
gene	complement(4202..4333)
CDS	/locus_tag="Bc4292"
	complement(4202..4333)
	/locus_tag="Bc4292"
	/codon_start=1
	/transl_table=1
	/product="EMG2 protein"
	/protein_id="AAP11205.1"
	/db_xref="GI:29897931"
	/translation="MQNINNILVGVDVYPVKAKIVQGAEPHVIGILFVASVAHLPRK"
gene	complement(4701..5513)
CDS	/locus_tag="Bc4293"
	complement(4701..5513)
	/locus_tag="Bc4293"
	/codon_start=1
	/transl_table=1
	/product="hypothetical Cytosolic Protein"
	/protein_id="AAP11206.1"
	/db_xref="GI:29897932"
	/translation="MDNKKIVVVSVDSETADLVRAAMGQFPAPDIRRVVYVEDTGTLEKEVISAKSNQALICETLVKQPMRQVLTVEAAKEGVADYDIIGPLDIEELTGQVPRYEGVVRUDEEYFKKIEAIEFAVKTIDDDRGDARGILKADIVLIGVSRSTKPLSQYLAHNRKLVANPLVLPVDPPEELYQVAKECFGLKITPDKLNHIRKERLKSGLSDGATYANLRIOBEIDHFBEVISRINCQVIDVSNKAIEETANIIVNAVQNRMP"
gene	complement(5543..6175)
CDS	/locus_tag="Bc4294"
	complement(5543..6175)
	/locus_tag="Bc4294"
	/codon_start=1
	/transl_table=1
	/product="CBS domain containing protein"
	/protein_id="AAP11207.1"
	/db_xref="GI:29897933"
	/translation="MIIELNKRHEHIIQIVKDHGPTGESIAAQLGLTRATLRPDLALFLMAGVLEAPRVGYFTYGTGGQLLSEAVYKIKVQDYQSRPVVIDKNVSVYDAICTMLEDVGTGLFVVDQMTLVGVVSKDILLRASLGKQDLTSLFVNIIMTRMPNIAACRREDSLVDIAELLERQIDAMPVVKDTQKQGLEVIGRITKNTITRAFNVLVNN"
gene	complement(6283..7029)
CDS	/locus_tag="Bc4295"
	complement(6283..7029)
	/locus_tag="Bc4295"
	/codon_start=1
	/transl_table=1
	/product="DNA repair protein recO"
	/protein_id="AAP11208.1"
	/db_xref="GI:29897934"
	/translation="MFQKVEGIVIRTTDYGTNKIIVTSRELKGVSAWAGSKKPKPSRLASVQLMTHGFLIQMGSLTQOGEIISTMKIEREDIFLTAVASFIVELTDKATBDKHNPLYFEMLYLHYMCEGVDPEVLISLYQTKMLPVLMRPFYDTCAICHOETD FVAFSREGGFLCSRHABQDPVRIPVGSAVHKLLRLLPFHDLHLRLGNVSVKDSKKQOMRLVLNTYYDEYCGIYILKSRRELEQLDKFQI"
gene	complement(7063..7206)
CDS	/locus_tag="Bc4296"
	complement(7063..7206)
	/locus_tag="Bc4296"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="AAP11209.1"
	/db_xref="GI:29897935"
	/translation="MLDFTWKFFSKTGSITETVLLLKEMEKDVNDMDQHEEELAHLDS PIS"
gene	complement(7383..8288)
CDS	/locus_tag="Bc4297"
	complement(7383..8288)
	/locus_tag="Bc4297"
	/codon_start=1

```
/transl_table=11
/product="GTP-binding protein"
/protein_id="AAP11210.1"
/db_xref="GI:29897936"
/translation="MNRKSGSGFVSIIGRNVGKSTFLNRIIGQKIAIMSDKPKQTR
NKIQGVYTFENDAQVIFDTPGHPKHLGDFMVKMAQTTLKEVDIVLFMVAVEGFG
RGEFFIEKLKTPQVFLVINKIDQLHPQLLELDQYRKLHFAETVPIFSAIDGNN
VDALIGTKLPEGPQVYPENQVTDHPERFIAELIREKVLHLTRFEEVPHSVAVID
AIOKEGAVYNATIVVERASOKGIIIGKQKMLKEVGRKARFDIEALLGSKVFLEY
WVKQKDRNKSQRDLGPREDEY"
/complement(8281..8679)
/locus_tag="BC4298"
/complement(8281..8679)
/EC_number="3.5.4.5"
/codon_start=1
/transl_table=11
/product="Cytidine deaminase"
/protein_id="AAP11211.1"
/db_xref="GI:29897937"
/translation="MNSKQILIEAEARKQAVVPYSKFQVGAALLTRGSKVYRGCVNE
NASTGLNCERTALFKAVSEGDFEVAIAVVADTKRPVPPCGACRQVMVELKQDTK
VYLSNLHGVDQVETTGELLGAPLAEDLHE"
/complement(8780..9133)
/locus_tag="BC4299"
/complement(8780..9133)
/locus_tag="BC4299"
/EC_number="2.7.1.107"
/codon_start=1
/transl_table=11
/product="Diacylglycerol kinase"
/protein_id="AAP11212.1"
/db_xref="GI:29897938"

Query Match      14.5%; Score 146.4; DB 1; Length 301332;
Best Local Similarity 55.4%; Pred. No. 1.3e-13;
Matches 310; Conservative 0; Mismatches 241; Indels 9; Gaps 1;

QY 25 AAAATTTTAAGTGAATTCGGCTTATTTGTTTGTATTTAAATGCACTGCAGCATGTGGAAT 84
DB 240517 AAAATTCAGTATTTTTCATAGTAGTCTTCTTATTCGCTGTGGATGCGGACAGCAAAA 240576

QY 85 AATAGTTCAAGTAACCAAGTAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCAGCAAA 144
DB 240577 GAGGAGAAAAGAAAACAAAGCGGACAATAAAATCAAGCTATAACAAATTAACACGCT 240636

QY 145 GAAGGTACTACGAAGTACTAAACACCCCTAAACGGTGTGTGTTCTTCAGTATTCATTT 204
DB 240637 GAAGGGAAACGAAGTTAGATAAACCCAGCAAAAAGATAGTTGTACTTCAGTGGGTATAT 240696

QY 205 GTTGATCGGTAGTGTCTTTAGATGTATAACCTGTGGGATAGCGATGATACAAAAA 264
DB 240697 TCTGAAGACTTATTAGCACTTGTGTTCAGCCAGTAGGATGCGACACATTAAGAATTAT 240756

QY 265 AATCGTATTATTAACCACTTAAGAGATAAAATGGAAATATACACTTCTGTAGGAACAGT 324
DB 240797 AATAATGGGTAATAACAAAACAAACCGAGTAAGATGTGTAGATGTAGGACACGT 240816

QY 325 AAGCAACCTAACCTTAGAAGAAATCAGTAACTTAAACCCAGATTTAATTTGCTGATAT 384
DB 240817 CAACACCAACCTTAGAAGAAATTAGCCGTTTAAACCCAGACTTAATTTATACAGCTTCA 240876

QY 385 AATAGACACAAAGGTATTTATAAGACTTAATAAATTTGCTCTCAGAT-----T 435
DB 240877 TTCGCGGTAAAGCAATTAATAATTAAGATTAAGACCAATTTGCACCAACAGTATGTTGAT 240936

QY 436 GAACGTAAGAGTTTCGATGAGATTAATAAGAAATATTCATGCTTTTAAACCAATTTCA 495
DB 240937 CCATCAACAAGCAATTAACATCACTTTCCTGAAATGACAGAAACATTTAAACAAATGCA 240996

QY 496 AAAGCTTTAGCTAAAGAAAGAGAGGTAAAAAAACGCTTAGAAGAACACGATAGAAAAAT 555
DB 240997 AAAGCAGTTTGAAGAAAGAGAGAGGTAAAAAAAGATTAGCTGATGATGAAGCATTC 241056
```

```
QY 556 GAAGAATATATAAAAGAAAT 575
DB 241057 GCTGATCAAAAGCAAAAT 241076

RESULT 30
AE017225 43
WPCOMMENT
Sequence split into 53 fragments LOCUS AE017225 Accession AE017225
Fragment Name Begin End
AE017225_00 1 110000
AE017225_01 100001 210000
AE017225_02 200001 310000
AE017225_03 300001 410000
AE017225_04 400001 510000
AE017225_05 500001 610000
AE017225_06 600001 710000
AE017225_07 700001 810000
AE017225_08 800001 910000
AE017225_09 900001 1010000
AE017225_10 100001 1110000
AE017225_11 1100001 1210000
AE017225_12 1200001 1310000
AE017225_13 1300001 1410000
AE017225_14 1400001 1510000
AE017225_15 1500001 1610000
AE017225_16 1600001 1710000
AE017225_17 1700001 1810000
AE017225_18 1800001 1910000
AE017225_19 1900001 2010000
AE017225_20 2000001 2110000
AE017225_21 2100001 2210000
AE017225_22 2200001 2310000
AE017225_23 2300001 2410000
AE017225_24 2400001 2510000
AE017225_25 2500001 2610000
AE017225_26 2600001 2710000
AE017225_27 2700001 2810000
AE017225_28 2800001 2910000
AE017225_29 2900001 3010000
AE017225_30 3000001 3110000
AE017225_31 3100001 3210000
AE017225_32 3200001 3310000
AE017225_33 3300001 3410000
AE017225_34 3400001 3510000
AE017225_35 3500001 3610000
AE017225_36 3600001 3710000
AE017225_37 3700001 3810000
AE017225_38 3800001 3910000
AE017225_39 3900001 4010000
AE017225_40 4000001 4110000
AE017225_41 4100001 4210000
AE017225_42 4200001 4310000
AE017225_43 4300001 4410000
AE017225_44 4400001 4510000
AE017225_45 4500001 4610000
AE017225_46 4600001 4710000
AE017225_47 4700001 4810000
AE017225_48 4800001 4910000
AE017225_49 4900001 5010000
AE017225_50 5000001 5110000
AE017225_51 5100001 5210000
AE017225_52 5200001 5228663
Continuation (44 of 53) of AE017225 from base 4300001 (AE017225 Bacillus anthracis str.
Query Match      14.2%; Score 143.2; DB 1; Length 110000;
Best Local Similarity 55.0%; Pred. No. 5.2e-13;
Matches 308; Conservative 0; Mismatches 243; Indels 9; Gaps 1;

QY 25 AAAATTTTAAGTGAATTCGGCTTATTTGTTTGTATTTAAATGCACTGCAGCATGTGGAAT 84
DB 33922 AAAATTCAGTATTTTTCATAGTAGTCTTCTTATTCGCTGTGGATGCGGACAGCAAAA 33981
```

QY	85	AATAGTTCAGTAACCTCAAGTAAGAGAGTCATCAAAAAGATGGAGTTGAAATCAAGCACGAA	144
Db	33982	GAGGAGAAAAAGAAACAAAACGCGACAAATAAAAATCAAGCTATAACAAATTAACACGCT	34041
QY	145	GAAGGTACTACGAAGTACCTAAACACCTAAACCGTGTGTGTCTTGAGTATTCATTT	204
Db	34042	GAAGGGGAAACGAAGTTAGATAAAACGCGAAAAAAGTAGTTGTACTTTGAATGGGTATAT	34101
QY	205	GTTCATGCTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATGATAACAAAAA	264
Db	34102	TCAGAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGATGGCAGACATTAAAGATTTAT	34161
QY	265	AATCGTATTATTAACACCTTAAGAGATAAATAATGGAATAATCACCTTCGTAGGAAACAGT	324
Db	34162	AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGACAGT	34221
QY	325	AAGCAACCTTAAGTAAAGAAATCAGTAAACCTTAAACCGAGATTTAAATTTATGCTGATAAT	384
Db	34222	CAACCAACCAACTTAGAAGAAATTAGCCGTTTAAACACAGATTTAAATTTATCACAGTTCA	34281
QY	385	AATAGACACAAAGTATTATTAAGACTTAAATAAAATTTGCTCCTACGAT-----T	435
Db	34282	TTCCGTGGTAAAGCAATTTAAATAATGAAATAGAACAAATTTGCACCAACAGTTATGTTGAT	34341
QY	436	GAAGTGAAGTTTCGATGGAGATTAATGAAATAATTTGATGCTTTTAAACAAATTTCA	495
Db	34342	CCATCAACAAAGCAATAACGATCACCTTTGCTGAAATGACAGAAACATTTTAAACAAATTTGCA	34401
QY	496	AAAGCTTTAGGTAAAGAGAGAGGTAAGAAACGCTTAGAAGACACGATAAGAAATTT	555
Db	34402	AAAGCAGTTGGAAAAAGAGAGAGGTAAGAAATTTAGCTGATATGATGATAAGACATTC	34461
QY	556	GAAGAATATATAAAAAAGAAAT	575
Db	34462	GCTGATGCATAAGCAAAAT	34481

RESULT 31
AE017334_43
WPCOMMENT

Sequence split into 53 fragments LOCUS AE017334 Accession AE017334

Fragment Name	Begin	End
AE017334_00	1	110000
AE017334_01	100001	210000
AE017334_02	200001	310000
AE017334_03	300001	410000
AE017334_04	400001	510000
AE017334_05	500001	610000
AE017334_06	600001	710000
AE017334_07	700001	810000
AE017334_08	800001	910000
AE017334_09	900001	1010000
AE017334_10	1000001	1110000
AE017334_11	1100001	1210000
AE017334_12	1200001	1310000
AE017334_13	1300001	1410000
AE017334_14	1400001	1510000
AE017334_15	1500001	1610000
AE017334_16	1600001	1710000
AE017334_17	1700001	1810000
AE017334_18	1800001	1910000
AE017334_19	1900001	2010000
AE017334_20	2000001	2110000
AE017334_21	2100001	2210000
AE017334_22	2200001	2310000
AE017334_23	2300001	2410000
AE017334_24	2400001	2510000
AE017334_25	2500001	2610000
AE017334_26	2600001	2710000
AE017334_27	2700001	2810000
AE017334_28	2800001	2910000
AE017334_29	2900001	3010000

AE017334_30	3000001	3110000
AE017334_31	3100001	3210000
AE017334_32	3200001	3310000
AE017334_33	3300001	3410000
AE017334_34	3400001	3510000
AE017334_35	3500001	3610000
AE017334_36	3600001	3710000
AE017334_37	3700001	3810000
AE017334_38	3800001	3910000
AE017334_39	3900001	4010000
AE017334_40	4000001	4110000
AE017334_41	4100001	4210000
AE017334_42	4200001	4310000
AE017334_43	4300001	4410000
AE017334_44	4400001	4510000
AE017334_45	4500001	4610000
AE017334_46	4600001	4710000
AE017334_47	4700001	4810000
AE017334_48	4800001	4910000
AE017334_49	4900001	5010000
AE017334_50	5000001	5110000
AE017334_51	5100001	5210000
AE017334_52	5200001	5227419

Continuation (44 of 53) of AE017334 from base 4300001 (AE017334 Bacillus anthracis str. Ames)

Query Match	14.2%;	Score 143.2;	DB 1;	Length 110000;	
Best Local Similarity	55.0%;	Pred. No. 5.2e-13;			
Matches 308;	Conservative	0;	Mismatches 243;	Indels 9;	Gaps 1;

QY	25	AAAAATTTAAAGTGTAAATTTGGCTTATTTGTTGTTTAAATTTGCAACTGCAGCATGTGGAAT	84
Db	33360	AAAAATTTCTCAGTATTTTTCATAGTAGTTTCTTCTATTCCTGTGTGGATGGGACAGCAAAA	33411
QY	85	AATAGTTCAGTAACCTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA	144
Db	33420	GAGGAGAAAAAGAAACAAAAGCGGACAAATAAAAATCAAGCTATAACAATTAACACGCT	33477
QY	145	GAAGGTACTACGAAAGTACCTAAACACCCCTAAACCGTGTGTGTTCTTGAGTATTCATTT	204
Db	33480	GAAGGGGAAACGAAAGTTAGATTAACCCAGCGAAAAAGTAGTTGTACTTCAATGGGTATAT	33533
QY	205	GTTCATGCGCTTAGTTGCTTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAA	264
Db	33540	TCAGAAAGACTTTATAGCACTTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAAATTA	33599
QY	265	AATCGTATTATTAACCACTTAAGAGATAAAATTTGGAATAACACTTCTGTAGGAACACAGT	324
Db	33600	AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT	33659
QY	325	AAGCAACCTTAAGAGAAATCAGTAAACCTTAAACACAGATTTAAATTTATGCTGATAAT	384
Db	33660	CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAACACAGATTTAAATTTATCACAGCTTCA	33711
QY	385	AATAGACACAAAGGTATTATTAAGAGCTTAAATAAAATTTGCTCCTACGAT-----T	435
Db	33720	TTCCGTGGTAAAGCAATTTAAAAATGAAATTTAGAACAAATTTGCCAACAGTTATGTTTGAT	33777
QY	436	GAAGTGAAGAGTTTCGATGGAGATTAATGAAATATTGATGCTTTTAAACAAATTTCA	495
Db	33780	CCATCAACCAAGCAATTAACGATCACTTTGCTGAAATGACAGAAACATTTTAAACAAATTTGCA	33833
QY	496	AAAGCTTTAGGTAAAGAGAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG	555
Db	33840	AAAGCAGTTGGTAAAGAGAGAGAGAGTAAAGAGTATTAGCTGATATGATTAAGCATTC	33899
QY	556	GAAGAATATATAAAAAAGAAAT	575
Db	33900	GCTGATGCAAAAGCAAAAAT	33919

RESULT 32
AE017038
LOCUS

AE017038 290029 bp DNA linear BCT 30-APR-2003

RESULT 32
AE017038
LOCUS

AE017038 290029 bp DNA linear BCT 30-APR-2003

GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,

GB:X59612, SP:P05"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP28211.1"
/db_xref="GI:30259006"
/translation="MSIRITQLVKSVYSPKQWALFRFQKIGKTYILMLCLITPIPT
FFYGSFTQDSVTVMNQIEKDLPDFKIENGELKADIEPIQKEEGDALFVDPNTDI
EYQNKTLGFLTKDKVWSIGNGQTYYSNDLGLTSLEKDLQDFISVFDSPYLLA
VIGLVVLQFLPFIPLGTTLLAFIGSAMSGORKLSYKQVMTLTAYSITIPIFFMIMD
LPKIVPQGSTPIYIAVLVILYLIKEVPKNEK"

gene

5799..6164

/locus_tag="BA4501"

CDS

5799..6164

/locus_tag="BA4501"

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AAP28212.1"

/db_xref="GI:30259007"

/translation="MKRLGIFLVLVLYGYFYDIKIGTLPLMLSSYKKTNAQTIKOE
STNWKONKAKKETDVTYKTIETVKTGETVLSITEQINKKKIPISEIKVDDFPQLNK
STSAIKQIGKSYKFLYQ"

gene

6329..7432

/gene="gcpB"

/locus_tag="BA4502"

CDS

6329..7432

/locus_tag="BA4502"

/note="identified by match to PFAM protein family HMM

PF02784"

/codon_start=1

/transl_table=11

/product="gcpE protein"

/protein_id="AAP28213.1"

/db_xref="GI:30259008"

/translation="MTHRTKTRPVKVNLTGGNNELIQSMTTTKTHDVEATVABIK
RLEAGCQVAVVADDERAANAADIKKQINIPVADIFPDYRLAKAIEGGIDKVI
NPGNIGRKHVAVVNAKERGIPIRIGVNAKSLERHLEKYGYPTADGMVESALHHI
KILBDLPHDIIVSMKASDVNLAIKAYKAAAFDYPPLHIGITBSGTLFAGTVKSAAG
LGAILNKIGNTLRISLSDPVEEVKVARLLKSPGLASNAATLISCTCGRIEIDL
SIANVEEYIITLQPIKVAIVGCAVNGFGAREADIGIAGAGEGLLFRKQGVVRKV
PEETNVEELKEIDVIAEMAAEREKEKETQEQ"

gene

complement(7513..7926)

/locus_tag="BA4503"

CDS

complement(7513..7926)

/locus_tag="BA4503"

/note="similar to GB:S67291, GB:S67292, GB:S67294,

GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,

GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,

GB:X59612, SP:P05230, PID:181942, PID:182559, PID:292022,

Query Match

Best Local Similarity 14.2%; Score 143.2; DB 1; Length 290029;

Matches 308; Conservative 55.0%; Pred. No. 4.1e-13;

Mismatches 243; Indels 9; Gaps 1;

QY 25 AAAATTTTAAAGTGAATGGCTTATTGTTGTTTAAATGCACTGCACGATGTGAAAT 84

DB 243302 AAAATTTCTAGTATTTTCATAGTAGTATTTTCTATTTCGCTGTGGATCGGACGACAAAA 243361

QY 85 AATAGTTCAAGTAACCAAGTAAAGATCATCAAAAAGATGGAGTTGAAATCAAGCACGAA 144

DB 243362 GAGGAGAAAAAGAAACAAACGCGACAATAAAATCAAGCTATAACCAATTAACACGCT 243421

QY 145 GAAGTACTACGAAGTACCTAAACACCTTAACCTGTGTTGTTCTTGGATTTCATTT 204

DB 243422 GAAGGGGAAACGAAGTTAGATAAACACGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 243481

QY 205 GTTGATGGCTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATACAAAAA 264

DB 243482 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGATGCGACACATTAAGAATTAT 243541

QY 265 AATCGTATTATTAAACCATTAAGAGATAAAATTTGGAAAAATACACTTCTGTAGGACACGT 324
DB 243542 AATAAATGGGTAATAACAAAAACAAACCCGAGTAAAGATGTTGTAGATGTCGGGACACGT 243601
QY 325 AAGCAACCTAACTTAGAAGAAATCAGTAAACCTTAAACCCAGATTTAAATTTATGCTGATAAT 384
DB 243602 CAACAACCAAACTTTAGAGAAATTTAGCGGTTTAAACCCAGATTTAAATTTATCAGACTTCA 243661
QY 385 AATAGACACAAAGGTATTTATAAAGACTTTAAATTTAAATTTGCTCTAGCAT-----T 435
DB 243662 TTCCGTGCTGAAGCAATTTAAAAATGAATTTAGAACAAATTTGCCAACAGTTATGTTTGAT 243721
QY 436 GAATCGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAACCAATTTCA 495
DB 243722 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAACCAATTTAAACCAATTTGCA 243781
QY 496 AAAGCTTTAGGTAAGAAGAAAGAGTAAAAAAACGCTTAGAAGAACACAGATAAGAAAAAT 555
DB 243782 AAAGCAGTTGCAAAAGAAAGAGAGGTAATAAAGTATTAGCTGATATGATTAAGCATTC 243841
QY 556 GAAGAATATAAAAAAGAAAT 575
DB 243842 GCTGATGCAAAAGCAAAAAAT 243861

RESULT 33

AE017355_43

WPCOMMENT

Sequence split into 53 fragments LOCUS AE017355 Accession AE017355

Fragment Name	Begin	End
AE017355_00	1	110000
AE017355_01	100001	210000
AE017355_02	200001	310000
AE017355_03	300001	410000
AE017355_04	400001	510000
AE017355_05	500001	610000
AE017355_06	600001	710000
AE017355_07	700001	810000
AE017355_08	800001	910000
AE017355_09	900001	1010000
AE017355_10	1000001	1110000
AE017355_11	1100001	1210000
AE017355_12	1200001	1310000
AE017355_13	1300001	1410000
AE017355_14	1400001	1510000
AE017355_15	1500001	1610000
AE017355_16	1600001	1710000
AE017355_17	1700001	1810000
AE017355_18	1800001	1910000
AE017355_19	1900001	2010000
AE017355_20	2000001	2110000
AE017355_21	2100001	2210000
AE017355_22	2200001	2310000
AE017355_23	2300001	2410000
AE017355_24	2400001	2510000
AE017355_25	2500001	2610000
AE017355_26	2600001	2710000
AE017355_27	2700001	2810000
AE017355_28	2800001	2910000
AE017355_29	2900001	3010000
AE017355_30	3000001	3110000
AE017355_31	3100001	3210000
AE017355_32	3200001	3310000
AE017355_33	3300001	3410000
AE017355_34	3400001	3510000
AE017355_35	3500001	3610000
AE017355_36	3600001	3710000
AE017355_37	3700001	3810000
AE017355_38	3800001	3910000
AE017355_39	3900001	4010000
AE017355_40	4000001	4110000
AE017355_41	4100001	4210000

1000001	1110000
BX571857_10	
BX571857_11	1210000
BX571857_12	1310000
BX571857_13	1410000
BX571857_14	1510000
BX571857_15	1610000
BX571857_16	1710000
BX571857_17	1810000
BX571857_18	1910000
BX571857_19	2010000
BX571857_20	2110000
BX571857_21	2210000
BX571857_22	2310000
BX571857_23	2400001
BX571857_24	2510000
BX571857_25	2600001
BX571857_26	2710000
BX571857_27	2700001
LOCUS	BX571857 2799802 bp DNA circular BCT 24-JUN-2004
DEFINITION	Staphylococcus aureus strain MSSA476, complete genome.
ACCESSION	BX571857
VERSION	BX571857.1 GI:49243355
KEYWORDS	complete genome.
SOURCE	Staphylococcus aureus subsp. aureus MSSA476
ORGANISM	Staphylococcus aureus subsp. aureus MSSA476
REFERENCE	Bacteria, Firmicutes; Bacillales; Staphylococcus.
AUTHORS	1 (bases 1 to 2799802)
	Holden,M.T.G., Fells,E.J., Lindsay,J.A., Peacock,S.J., Day,N.P.J.,
	Enright,M.C., Foster,T.J., Moore,C.E., Hurst,D., Atkin,R.,
	Barron,A., Bason,N., Bentley,S.D., Chillingworth,C.,
	Chillingworth,T., Church,C., Clark,L., Corton,C., Cronin,A.,
	Doggett,J., Dowd,L., Fellwell,T., Hance,Z., Harris,B., Hauser,H.,
	Hooley,S., Jagels,K., James,K.D., Lennard,N., Line,A., Mayes,R.,
	Moule,S., Mungall,K., Ormond,D., Quail,M.A., Rabinowitch,B.,
	Rutherford,K., Sanders,M., Sharp,S., Simmonds,M., Stevens,K.,
	Whitehead,S., Barrett,B.G., Spratt,B.G. and Parkhill,J.
TITLE	Complete genomes of two clinical Staphylococcus aureus strains:
	evidence for the rapid evolution of virulence and drug resistance
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (26), 9786-9791 (2004)
REFERENCE	2 (bases 1 to 2799802)
AUTHORS	Holden,M.T.G.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUN-2004) Submitted on behalf of the Pathogen
	Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
	Hinxton, Cambridge CB10 1SA, E-mail: ml@sanger.ac.uk
FEATURES	Location/Qualifiers
source	1..2799802
	/organism="Staphylococcus aureus subsp. aureus MSSA476"
	/mol_type="genomic DNA"
	/strain="MSSA476"
	/db_xref="taxon:282459"
	517..1878
gene	/locus_tag="SAS0001"
	517..1878
CDS	/locus_tag="SAS0001"
	/note="Ortholog of S. aureus MRSA252 (BX571856) SAR0001"
	/codon_start=1
	/transl_table=11
	/product="chromosomal replication initiator protein Dnaa"
	/protein_id="CAG41773.1"
	/db_xref="GI:49243356"
	/translation="MSKEKEVLEIAEQEKLNAVSVSTFLKDTLYTKDGEATVLA
	SIFPNANLWQQVAAEIIQALIFDVGVGEVPHFITTELFANYSNNETATPKETTPRS
	ETEDHNLVAGREFNAETFDFTVIGCNRPFFHAASLAAVEAPKAYNPLFYGGVG
	KTHLMHIAIHHVLDNNPDARVITYTSKEFTNEFKGIRDNDEGAFAFRERYNIDVLL
	DDIQFONKYQTOEEFYTNELHNNKQIVISSDRPPEKLAQLEGRRLRSYFEGWGLI
	DTFPDYTEMALLOKIESEKLDIPEALNYIANQTSNIRELEGALTRLLAYSQL
	GKPTTETAEALKDIIQAPEKSKITTIQDIQIKVQYVNVPIEDFSAKRTKSIAPY
	QIANYLSREITDLSFKIGEGFGDRHTTVIHAHEKISKOLKEDPIFKQEVENLEKE
	RNV"
misc_feature	862..1803
	/locus_tag="SAS0001"
	/note="Pfam match to entry PF00308 bac_dnaA Bacterial

```

dnaa protein, score 681.8, E-value 3.5e-202"
982. .1005
/locus_tag="SAS0001"
/notes="PS00017 ATP/GTP-binding site motif A (P-loop)."
1744. .1803
/locus_tag="SAS0001"
/notes="PS01008 DnaA protein signature."
2156. .3289
/locus_tag="SAS0002"
2156. .3289
/locus_tag="SAS0002"
/EC_number="2.7.7"
/notes="Ortholog of S. aureus MRSA252 (BX571856) SAR0002"
/transl_table=11
/product="DNA polymerase III, beta chain"
/protein_id="CAG41774.1"
/db_xref="GI:49243357"
/translation="MMETIKRDYFETQLNDTLKAIISPRFTLPILTGIKIDAKEHEVI
LTGDSRISITITPKTVGDGDIYNISETGVSVLPGRFVDIIKKLPQKOVKLSTNEQ
FQTLITSGHSFNLSGDDPDQYPLLPQVSRDDAIQLSVKVLKNIAQTNFAVSTSETR
PVLTVGNLMIQENELICTATDSHRLAVRKLQLEDVSENKNVIFPKALAEINKFMSDN
BEDIDFASQNVLFVGNVNFISRLLEGHYPTDTRLFPFNEYIEKLSIDNGEFPYHAI
RASLLAREGNNVTKLSTGDDVVELSTSPRIGTVKEVDANDVEGGLSKISFNSKYM
MDALKAINDRSEVEFFGTMKPFILKPKGDDSVTLILPIRTY"
2159. .2539
/locus_tag="SAS0002"
/notes="Pfam match to entry PF00712 DNA_pol3_beta, DNA
polymerase III beta subunit, N-terminal domain, score
224.5, E-value 1.6e-64"
2564. .2908
/locus_tag="SAS0002"
/notes="Pfam match to entry PF02767 DNA_pol3_beta_2, DNA
polymerase III beta subunit, central domain, score 222.4,
E-value 7e-64"
2912. .3280
/locus_tag="SAS0002"
/notes="Pfam match to entry PF02768 DNA_pol3_beta_3, DNA
polymerase III beta subunit, C-terminal domain, score
209.8, E-value 4.4e-60"
3670. .3915
/locus_tag="SAS0003"
3670. .3915
/locus_tag="SAS0003"
/notes="Ortholog of S. aureus MRSA252 (BX571856) SAR0003"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAG41775.1"
/db_xref="GI:49243358"
/translation="MILIVQVSVVEGDINLQFKTEGIIESGGQAKWFLOQDVVLIN
GVRETRGKKLEHQDRIDIPELPEDAGSFLIIHQEQ"
3912. .5024
/locus_tag="SAS0004"
3912. .5024
/locus_tag="SAS0004"
/notes="Ortholog of S. aureus MRSA252 (BX571856) SAR0004"
/codon_start=1
/transl_table=11
/product="DNA replication and repair protein RecF"
/protein_id="CAG41776.1"
/db_xref="GI:49243359"
/translation="MKNLTLQENVRNRYDEVTLKCHPDVNVILIGENAGKTNLLESYI
TLAKSHRTSNDKELIRFNADYAKIEGELSYRHGTMLPTMFTTKGQKQVKNHLEQS
RLTYIGHNLVLPFAPEDNI VKGSPQIRRRFIDMELGQISAVYLNDAQYRIKQK
NNYLKQGLQKDLTMLEVLNQQFAEYAMKVTOKRAHFIOELSLAKPIAHITNDK
BALSNLYLPSLKFDYAONEAARLEIMSIISDNMREKERGISLFGPHRDDISFDVNG
MDQTYGSGQQRQRTALSKLAELIEMNIEVEYPIILDDVLSELDDSRQTHLLSTI
HQKQVTFVTVTSVDIHEIMNNAKLRYINQGEILK"
3912. .4526
/locus_tag="SAS0004"
/notes="Pfam match to entry PF02463 SMC N, RecF/RecN/SMC N
terminal domain, score -21.4, E-value 2.5e-05"

```

```

misc_feature
3999. .4022
/locus_tag="SAS0004"
/notes="PS00017 ATP/GTP-binding site motif A (P-loop)."
4242. .4319
/locus_tag="SAS0004"
/notes="PS00617 RecF protein signature 1."
4848. .4901
/locus_tag="SAS0004"
/notes="PS00618 RecF protein signature 2."
5037. .6968
/locus_tag="SAS0005"
5037. .6968
/locus_tag="SAS0005"
/EC_number="5.99.1.3"
/notes="Ortholog of S. aureus MRSA252 (BX571856) SAR0005"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="CAG41777.1"
/db_xref="GI:49243360"
/translation="MTALSDVNTDNYGACIQVLEGLBAVRKPKMYIGSTSERGLH
HLVWEIVDNIIDEALAGYANQIEVIEKDNWIKVTDNGRGI PVDIQEMKRGPAVEVIL
HLVHAGKPGGGYKVGSGHGVSSVYNALSQDLVEVYVHRNETIYHOAYKKGVPQFD
LKEVYTDKGTIVIRFKADGEIETETVYNYETIQQRIRBLAFLNKGITLDRDERE
ENVREDSYHYEGGIKSVLLENKEPIHDBPIYHOSKDDIEVEIAIQNSGYATNL
LTVANNHTYEGGTHEDGFKRALTRVLNSGLSKIMKERKDLRSGDTRGEMTAIIS
IKGDPQFEGQTKIKLGNSEVRQVVDKLPSEHFERFLYENPQVARTVEKIGMAARAR
VAAKAREVTRRSALDASLPGLKADCSSPEECIFLVEGDSAGSTKMSGRSRT
QAILPLRGKILNVKARLDRLNNREIRQMITAFGTGIGGDFDLAKARYHKIVIMTDA
DVCAHIRTLLLPFYFRMRPLTEAGYVYTAQPLKLTQCKYKQYVYNDRLELDKLS
ELNPTPKWSIARYKGLGEMNADQLWETTMNPEHRALLQVLEDAIEADQTFPEMLMGDV
VENERQFIEDNAVYANLDF"
5148. .5579
/locus_tag="SAS0005"
/notes="Pfam match to entry PF02518 HATPase_C, Histidine
kinase-, DNA gyrase B-, and HSP90-like ATPase, score
103.9, E-value 3.4e-28"
5538. .5558
/locus_tag="SAS0005"
/notes="PS00154 Ei-E2 ATPases phosphorylation site."
5727. .6239

```

```

Query Match      12.5%; Score 125.6; DB 1; Length 110000;
Best Local Similarity 50.9%; Pred. No. 2.4e-10;
Matches 413; Conservative 0; Mismatches 379; Indels 20; Gaps 4;

QY 27 AATTTAAGTGTAAATGGCTTAATTTGTTTAAATGCAACTGCAGCATGTGGAATAA 86
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104242 AATTTAAATGCTTGTGTTACGCTTGCTTCTCTACTTGTTTTAGCAGCATGTAGTGGAA 104183

QY 87 TAGTTCAGTAACTCAAGTAAAGACTCATCAAAGATGGAGTTGAAATCAAGCAGAGA 146
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104182 TTCAAATAAACAATCATCTGTATACAAAGATAAGGAACAACATTCATTAACATGCAAT 104123

QY 147 AGGTACTACGAAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGAGTATTCATTTGT 206
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104122 GGGTACAACTGAAATTAAGGAAACCAACCGCTGTTGTTACGTATATATCAAGTGCAC 104063

QY 207 TGATCGGTAGTTCCTTTAGATGTTAAACCTGTTGGGATACGGGATGATACAAAAAAAA 266
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104062 TGACGTGCTGTATCTTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAA 104003

QY 267 TCGTATTATTAACCATTAAGAGATAAAATGGAAATACACTTCTGTAGGAACACGTAA 326
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 104002 ACCGAAATTCGAATACATAAAAAATGATTTAAAAAGATACATAAGATTGTAGTCAAGAAC 103943

QY 327 GCAACCTAACCTAGAGAAATCAGTAAACTTAAACACAGATTAAATTTGCTGTAATAA 386
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 103942 TGCACCTAACTTAGGAAATCTCTAAATTAACCCGACTTAATTTGCGGTCAAAAGT 103883

QY 387 TAGACACAAAGGTATTTATAAGACTTAATAATAAATTTGCTCTACGATTGAACATGAAGA 446
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 103882 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAAATCGCACC-----AACAG 103836

```


CDS

10893..12407
 /gene="huth"
 /note="ORFID:MW0008"
 /codon_start=1
 /transl_table=1
 /product="histidine ammonia-lyase"
 /protein_id="BAB93873.1"
 /db_xref="GI:21203172"
 /translation="MTIYLDGETITIEDIKSFLOQSKIEIIDALERVKKSRAVVER
 IENEEYTGITGFGFSDVRIPTQYNLQVNLIRSHACGLGEPFSEKVALVMML
 RLNTLKHGSGATDLRLNQLQFFNIERIPIPOQSGLSAGSDGLAPLSHLALIGEG
 KVALRGESKDDVLRNQLQAKEGALLINGLNGAQTQAMTAQGVISVISEDILGYQS
 EWTAAALHQSLANGIIDAVRHVDHVRVFOEINVAARMRDLEGSTLTTCROAEIRVQD
 AYTLRCIPQHGASFOVNYKQLEFEMNAANDNPLIPEANETFFVSGNEFRGQPI
 AFADHLKGVSELANYSERLERLVNPLQMLDLPFLSPFGPQSGAMINQWIAAASL
 VSENKTLGSPASVDSITYSSANQEDHVSNTTAARHGQYIENARRVIAECVIALQAA
 ELKVGLEASPTRKRYBEFRSIVPSITHDROGKIDIEAVQYLKQSIYQTACH"
 12786..14072
 /gene="gerS"
 /note="ORFID:MW0009"
 /codon_start=1
 /transl_table=1
 /product="eeryl-tRNA synthetase"
 /protein_id="BAB93874.1"
 /db_xref="GI:21203173"
 /translation="MELDIRFNEPFTVYKSKIELRGDDPKVDVILELDEQRKLISA
 TEEMKARNKVSSEIALKRNKNADDDVIAEMRTLGDDIKEKDSQLNEDNMTGILC
 RIPNLIISDDVQGSDEBNVKKWTPREFSPKAWHDIVBELKMDPDKRAKVSQ
 ARFVYLPMQALSERALNMVMTKHTTQHGVTMMVPOLVNAQDMYGTGQLPKTEEDL
 FKVEKEGLYIPIAEVLPTNRYNEIIQPGLVPEKFTGQSACFRSEAGSAGDRGLI
 RLHQFDKVMVRFQPSDQWNALEEMITNAEILEUGLPYRRVILCTGDIIGFSASKT
 YDLEVMFVSYNDYKEIISCSNCTDQFARRNIRFKRDKAAPKELAHNTLNGSLAVGR
 FFAATVENVQNEQDGTWTIPEALVPFMGGKTOISKVPK"
 14724..15419

gene

CDS

Query Match 12.5%; Score 125.6; DB 1; Length 290250;
 Best Local Similarity 50.9%; Pred. No. 1.9e-10;
 Matches 413; Conservative 0; Mismatches 379; Indels 20; Gaps 4;

QY 27 AATTAAATGCTAATGGCTATTGTTGTTTATTTAATTCGCACTGCAGCATGTGGAATAA 86
 DB 105536 AATTAAATGCTGTTGTTGACGTTGCTTTCTACTGTTTATGACGAGATGTAGTGGNA 105477

QY 87 TAGTTCAGTAACCTCAAGTAAGAGTCAATAAAGATCGAGTTGAAATCAAGACGAGA 146
 DB 105476 TTCAATAAACAATCATCTGATAACAAAGATAAGGAAACAACCTTCAATTAAACATGCAAT 105417

QY 147 AGGTACTACGAAAGTACCTAACACCCCTAAACGTTGTTGTTCTTGAGTATTCATTTGT 206
 DB 105416 GGGTACAACTCAAAATTAAGGGAAACCAAAGCGTTGTTGACGTATATCAAGGTGCAC 105357

QY 207 TGATGCGTTAGTGTGCTTTAGATGTTTAAACCTGTTTGGGATAGCGGATGATAACAAAAA 266
 DB 105356 TGACGTGCTGTATCTTTAGGTGTTAAACCTGTAGTGTCTGTAGATCATGACACAAA 105297

QY 267 TCGTATTATTAACCATTAAGAGATAAAATTGGAATAATACACTCTCTGAGGAACACGTAA 326
 DB 105296 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATACTAAGATTGTAGGTCAAGAACC 105237

QY 327 GCAACCTAACTTAGAGAAATCAGTAACCTTAACACAGATTTAAATTTGCTGATAATAA 386
 DB 105236 TGCACCTAACTTAGAGGAAATCTCTAAATTAATAAACCGGACTTAATTTGCGGTCAAAAGT 105177

QY 387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAAAATTTGCTCTACGATTTGAACCGAAAG 446
 DB 105176 TAGAAATGAAAAAGTTTACGATCAATATCTAAATCGACC-----AACAG 105130

QY 447 TTTTCGATGAGATTATAATGAAAT-ATTGATGCTTTTAAACAAATTTCAAAGCTTTAG 505

DB 105129 TTTCTACTGATACAGTCTTTCAAATTCAAAGATACAACTAAGTAAATGCGGAAGCTTTAG 105070

Qy 818 CTAACGACCTTCTACTAAAGAACTAGAAAAGATCCTGTATGGAAGAAATTTAAACGCTG 877
Db 361 ACAAAA-----TAGAACGAAATTCATGTATCCTCGAGTTTGGAAATCATTTAAAGCTG 414
Qy 878 TGAATAATCAACGTTGTGATATTTTAGACCGTGACTTATGGCGCAAGATCACTGCTGTTTAA 937
Db 415 TGAAGATAACAAGTTTATGACGTTGACCGAATTAAGTGGTTGAAATCAAGGGGATTA 474
Qy 938 TTCTCTCAGAAAGATGCGCAAGAACTTGTGTGAATTTATCTAAGAAAGATAGTAAAAAG 997
Db 475 TCGCAAGTGAAATGATGCGCAGAGATTTAGAAAAAATTTGCAAAAAAGCAAAATAAAAAT 534
Qy 998 ATA 1000
Db 535 ACA 537
RESULT 39
AE017355_06/c
WPCOMMENT
Sequence split into 53 fragments LOCUS AE017355 Accession AE017355
Fragment Name Begin End
AE017355_00 1 110000
AE017355_01 100001 210000
AE017355_02 200001 310000
AE017355_03 300001 410000
AE017355_04 400001 510000
AE017355_05 500001 610000
AE017355_06 600001 710000
AE017355_07 700001 810000
AE017355_08 800001 910000
AE017355_09 900001 1010000
AE017355_10 1000001 1110000
AE017355_11 1100001 1210000
AE017355_12 1200001 1310000
AE017355_13 1300001 1410000
AE017355_14 1400001 1510000
AE017355_15 1500001 1610000
AE017355_16 1600001 1710000
AE017355_17 1700001 1810000
AE017355_18 1800001 1910000
AE017355_19 1900001 2010000
AE017355_20 2000001 2110000
AE017355_21 2100001 2210000
AE017355_22 2200001 2310000
AE017355_23 2300001 2410000
AE017355_24 2400001 2510000
AE017355_25 2500001 2610000
AE017355_26 2600001 2710000
AE017355_27 2700001 2810000
AE017355_28 2800001 2910000
AE017355_29 2900001 3010000
AE017355_30 3000001 3110000
AE017355_31 3100001 3210000
AE017355_32 3200001 3310000
AE017355_33 3300001 3410000
AE017355_34 3400001 3510000
AE017355_35 3500001 3610000
AE017355_36 3600001 3710000
AE017355_37 3700001 3810000
AE017355_38 3800001 3910000
AE017355_39 3900001 4010000
AE017355_40 4000001 4110000
AE017355_41 4100001 4210000
AE017355_42 4200001 4310000
AE017355_43 4300001 4410000
AE017355_44 4400001 4510000
AE017355_45 4500001 4610000
AE017355_46 4600001 4710000
AE017355_47 4700001 4810000
AE017355_48 4800001 4910000
AE017355_49 4900001 5010000
AE017355_50 5000001 5110000
AE017355_51 5100001 5210000
AE017355_52 5200001 537682
Continuation (7 of 53) of AE017355 from base 600001 (AE017355 Bacillus thuringiensis ser

Query Match 12.4%; Score 124.6; DB 1; Length 110000;
Best Local Similarity 48.2%; Pred. No. 3.5e-10;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2005, 23:38:01 ; Search time 644 Seconds
(without alignments)
9265.682 Million cell updates/sec

Title: US-10-724-972a-2580
Perfect score: 1008
Sequence: 1 ggaggggaatcagtgagagg.....gtaaaaagataataagtaa 1008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1008	100.0	1008	13	ADS03285 Staphyloc
2	993	98.5	993	8	ACA47084 Prokaryot
3	569	56.4	2981	8	ACC48532 Staphyloc
4	569	56.2	984	8	ACA19774 Prokaryot
5	564.4	56.0	1014	4	AAS4519 Staphyloc
6	557.4	55.3	1044	8	ACF74986 Staphyloc
7	531.8	52.8	927	4	AAS1822 Staphyloc
8	336.2	33.4	2247	2	AAT80398 Staphyloc
9	336.2	33.4	2247	2	AAT83786 DNA encod
10	309.8	30.7	525	8	ACA47237 Prokaryot
11	269.2	26.7	2957	8	ACC48534 Staphyloc
12	265.6	26.3	960	4	AAS4387 Staphyloc
13	264.6	26.2	927	4	AAS1612 Staphyloc
14	263.8	26.2	975	8	ACF74415 Staphyloc
15	263.2	26.1	957	8	ABT15015 Pathogen
16	263.2	26.1	960	8	ACA20079 Prokaryot
17	190.8	18.9	5718	12	ADF30765 Bacillus
18	175.4	17.4	668	2	AAV74666 Staphyloc
19	162.2	16.1	242	2	AAV76867 Staphyloc
20	158.2	15.7	193	6	ABN92560 Staphyloc

21	156.4	15.5	897	8	ACA42836	ACA42836 Prokaryot
22	149.4	14.8	801	6	ABK74775	ABK74775 Bacillus
23	143.2	14.2	972	8	ACA21545	ACA21545 Prokaryot
24	128	12.7	321	4	AAS49530	AAS49530 Staphyloc
25	128	12.7	321	8	ACA16735	ACA16735 Prokaryot
26	125.4	12.4	2115	2	AAV74915	AAV74915 Staphyloc
27	124	12.3	990	8	ABT15032	ABT15032 Pathogen
28	124	12.3	999	2	AZ22850	AZ22850 Staphyloc
29	124	12.3	999	2	AZ22850	AZ22850 Staphyloc
30	124	12.3	999	10	ADF43555	ADF43555 Staphyloc
31	124	12.3	2940	8	ACC48531	ACC48531 Staphyloc
32	124	12.3	3775	2	AAV74549	AAV74549 Staphyloc
33	113.2	11.2	796	8	ACA22139	ACA22139 Prokaryot
34	110.2	10.9	861	8	ACF74610	ACF74610 Staphyloc
35	109.4	10.9	3046	4	AAH54418	AAH54418 S. epider
36	103.2	10.2	900	8	ACA32000	ACA32000 Prokaryot
37	96.8	9.6	909	10	ACF70890	ACF70890 Photorhab
38	96.8	9.6	25860	11	ADR20887	ADR20887 Photorhab
39	96.8	9.6	110000	10	ACF67367	Continuation (40 o
40	96.8	9.6	110000	10	ACF65388	Continuation (9 of
41	96.2	9.5	1130	2	AAT83923	AAT83923 DNA encod
42	95.2	9.4	909	4	AAS52735	AAS52735 E. coli D
43	95.2	9.4	909	8	ACA32751	ACA32751 Prokaryot
44	95.2	9.4	2209	5	AAS71379	AAS71379 DNA encod
45	95.2	9.4	2799	5	AAS77335	AAS77335 DNA encod

ALIGNMENTS

RESULT 1
ADS03285
ID ADS03285 standard; DNA; 1008 BP.

AC ADS03285;

DT 04-NOV-2004 (first entry)

DE Staphylococcus epidermis polynucleotide seqid 2580.

KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system; gene; ds.

OS Staphylococcus epidermidis.

XX US2004147734-A1.

XX 29-JUL-2004.

XX 01-DEC-2003; 2003US-00724972.

XX 08-NOV-1997; 97US-0064964P.

XX 13-AUG-1998; 98US-00134001.

XX 29-NOV-1999; 99US-00450969.

XX (DOUC/) DOUCETTE-STAMM L.

XX (BUSH/) BUSH D.

XX Doucette-Stamm L, Bush D;

XX WPI; 2004-580138/56.

XX P-PSDB; ADS07057.

XX New isolated polypeptide and encoding nucleic acid derived from

XX Staphylococcus epidermidis, useful for diagnosing, preventing and/or

XX treating an S. epidermidis bacterial infection.

XX Claim 5; SEQ ID NO 2580; 741pp; English.

XX The invention describes an isolated nucleic acid comprising a nucleotide

XX sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:

XX 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any

59-67

CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC given in the specification. Also described are: a recombinant expression
CC vector; a cell comprising a recombinant expression vector of (1);
CC producing an S. epidermidis polypeptide; an isolated nucleic acid
CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection, comprising a nucleic acid cited above and a carrier; treating
CC a subject for S. epidermidis infection; a recombinant or substantially
CC pure preparation of an S. epidermidis polypeptide or its fragment; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection; detecting the presence of a Staphylococcus nucleic acid in a
CC sample; a computer readable medium having recorded in it the nucleotide
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC system for identifying fragments of the Staphylococcus genome of
CC commercial importance; a computer based system for identifying fragments
CC of the Staphylococcus plasmids of commercial importance; identifying
CC commercially important nucleic acid fragments of the Staphylococcus
CC genome and/or plasmids; and identifying an expression modulating fragment
CC of the Staphylococcus genome and/or plasmids. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of an Staphylococcal epidermidis bacterial
CC infection. This sequence encodes a S. epidermis protein of the invention.
XX
SQ Sequence 1008 BP; 404 A; 129 C; 188 G; 287 T; 0 U; 0 Other;

Query Match 100.0%; Score 1008; DB 13; Length 1008;
Beat Local Similarity 100.0%; Pred. No. 3.3e-175;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAGTGGAAATCAGTCAGAGGTTTAAAAATTTTAACTGTAATTTGGCTTTTGTGTTTAA 60
DB 1 GGAGTGGAAATCAGTCAGAGGTTTAAAAATTTTAACTGTAATTTGGCTTTTGTGTTTAA 60
QY 61 ATTGCAACTGACGATGTGGAATTAATAGTTCAAGTAATCACTCAAGTAAAGAGTCATCAAAA 120
DB 61 ATTGCAACTGACGATGTGGAATTAATAGTTCAAGTAATCACTCAAGTAAAGAGTCATCAAAA 120
QY 121 GATGAGTTGAAATCAAGCAGCAAGAGGTAAGTACTAGCAAGTAACTCAACACCCCTAAACGT 180
DB 121 GATGAGTTGAAATCAAGCAGCAAGAGGTAAGTACTAGCAAGTAACTCAACACCCCTAAACGT 180
QY 181 GTTGTGTTCTTGTAGTATTCATTTGTTGATGCTTTAGTGTGTTTGTGTTTAAACCTGTT 240
DB 181 GTTGTGTTCTTGTAGTATTCATTTGTTGATGCTTTAGTGTGTTTGTGTTTAAACCTGTT 240
QY 241 GGGATAGCGGATGATAACCAATTAATCGTATTTTAAACCACTTAAGAGATAAAATCGA 300
DB 241 GGGATAGCGGATGATAACCAATTAATCGTATTTTAAACCACTTAAGAGATAAAATCGA 300
QY 301 AAATACACTTCTGTAGGACACGTAAAGCACTTAAGTAAAGTAACTCAAGTAACTTAA 360
DB 301 AAATACACTTCTGTAGGACACGTAAAGCACTTAAGTAAAGTAACTCAAGTAACTTAA 360
QY 361 CCAGATTTAAATTTTGTGATTAATTAAGACAAAGGTTATTTATAAGACTTAAATAAA 420
DB 361 CCAGATTTAAATTTTGTGATTAATTAAGACAAAGGTTATTTATAAGACTTAAATAAA 420
QY 421 ATTGCTCTCTACGATTAAGTAAAGTTCGATGAGATTTAATGAAATATTCATGCT 480
DB 421 ATTGCTCTCTACGATTAAGTAAAGTTCGATGAGATTTAATGAAATATTCATGCT 480
QY 481 TTTAAACCAATTTCAAAAGCTTTAGTAAAGAAAGAGGTTAAACCGCTTAGAAGAA 540
DB 481 TTTAAACCAATTTCAAAAGCTTTAGTAAAGAAAGAGGTTAAACCGCTTAGAAGAA 540
QY 541 CACGATAAGAAATTTGAAGAAATTAATAAGAAATTAATGATGATAAAATCAAAAGTGA 600
DB 541 CACGATAAGAAATTTGAAGAAATTAATAAGAAATTAATGATGATAAAATCAAAAGTGA 600
QY 601 TTGCTGTCAGTACGCTGAATCAGGTTTGTGCTTCATCCCAAGCAACTCTTATGTTGT 660
DB 601 TTGCTGTCAGTACGCTGAATCAGGTTTGTGCTTCATCCCAAGCAACTCTTATGTTGT 660

QY 661 CAATTCCTCACTCACTAGGTTTAAAGAAAGCATTAAGTGTATGATGTTTAAAGGTTTA 720
DB 661 CAATTCCTCACTCACTAGGTTTAAAGAAAGCATTAAGTGTATGATGTTTAAAGGTTTA 720
QY 721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAAT 780
DB 721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAAT 780
QY 781 CCTGAGGCTATGTTTCAATATGACAAACAAAGCAAGTCTTAACGAACTTCACTAAAGAA 840
DB 781 CCTGAGGCTATGTTTCAATATGACAAACAAAGCAAGTCTTAACGAACTTCACTAAAGAA 840
QY 841 CTAGAAAGGATCCTGTATGGAAGAAATTAACCCCTGTGAAAGAAATCAACGTTGTGATATT 900
DB 841 CTAGAAAGGATCCTGTATGGAAGAAATTAACCCCTGTGAAAGAAATCAACGTTGTGATATT 900
QY 901 TTAGACCGTGACTTATGGGCAAGATCAGTGGTTTAAATTTCTCAGAGAAATATGCAAAA 960
DB 901 TTAGACCGTGACTTATGGGCAAGATCAGTGGTTTAAATTTCTCAGAGAAATATGCAAAA 960
QY 961 GAACCTGTTGATTTATCTAAGAAAGATAGTAAAGAAAGATATAAGTAA 1008
DB 961 GAACCTGTTGATTTATCTAAGAAAGATAGTAAAGAAAGATATAAGTAA 1008

RESULT 2

ACAA47084
ID ACA47084 standard; DNA; 993 BP.
XX ACA47084;
XX
XX 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #28741.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
XX Staphylococcus epidermidis.
OS
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Travick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU43214.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 34954; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the

antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 98.5%; Score 993; DB 8; Length 993; Best Local Similarity 100.0%; Pred. No. 1.9e-172; Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

13 GTGAGAGGTTTAAAAATTTAAAGTAAATGGCTTATTTGTTTAAATGCACTGCA 72
1 GTGAGAGGTTTAAAAATTTAAAGTAAATGGCTTATTTGTTTAAATGCACTGCA 60
73 GCATGTGGAATAATAGTTCACAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTGA 132
61 GCATGTGGAATAATAGTTCACAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTGA 120
133 ATCAAGCAGCAAGAGGTACTAGCAAGTACCTAAACACCCCTAAACGTTGTTGTTCT 192
121 ATCAAGCAGCAAGAGGTACTAGCAAGTACCTAAACACCCCTAAACGTTGTTGTTCT 180
193 GAGTATTCATTTGTTGATGGTGTAGTTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA 252
181 GAGTATTCATTTGTTGATGGTGTAGTTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA 240
253 GATAACAAAAAATCGTATTTATTAACCAATTAAGAGATAAAATTCGAAAAATACACTTCT 312
241 GATAACAAAAAATCGTATTTATTAACCAATTAAGAGATAAAATTCGAAAAATACACTTCT 300
313 GTAGGAACAGTAAGCAACCTTAAGAGATAAAATTCGAAAAATACACTTCT 372
301 GTAGGAACAGTAAGCAACCTTAAGAGATAAAATTCGAAAAATACACTTCT 360
373 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 432
361 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 420
433 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 492
421 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 480
493 TCAAAAGCTTTAGTAAAGAGAGAGGTAAAGAGAGGTAAAGAGAGGTAAAGAGAGGTAA 552
481 TCAAAAGCTTTAGTAAAGAGAGAGGTAAAGAGAGGTAAAGAGAGGTAAAGAGAGGTAA 540
553 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 612
541 ATTGCTGATTAATAGACCAAGGTATTTAAAGACCTTAATAAATTCGCTCTACG 600
613 GCTGCTAAATCAGTGTGTTGCTTCATCCCAAGCAACTTATGTTGTTGTTGTTGTTGTTG 672

601 GCTCTAAATCAGGTTTGGCTTCATCCAAAGCAACTTATGTTGGTCAATTCCTAAGT 660
673 CAATAGGTTTTTAAAGAGCAATTAAGTGCATGATGTTACTAAAGGTTTAAAGTATCTT 732
661 CAATAGGTTTTTAAAGAGCAATTAAGTGCATGATGTTACTAAAGGTTTAAAGTATCTT 720
733 AAAGACCTTACTTACAAATGAACACTGAAACTTTTCTCAAGTGAATCCTGAGCGTATG 792
721 AAAGACCTTACTTACAAATGAACACTGAAACTTTTCTCAAGTGAATCCTGAGCGTATG 780
793 TTCTAATGACAAACAAAGCAAGTTCTAACCAACCTTCACTAAAAAGAACTAGAAAAAGAT 852
781 TTCTAATGACAAACAAAGCAAGTTCTAACCAACCTTCACTAAAAAGAACTAGAAAAAGAT 840
853 CCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTAGACCGTAC 912
841 CCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATATTTTAGACCGTAC 900
913 TTATGGGCAAGATCACTGGTGTAAATTTCTTCAGAGAAATGGCAAAAGAACTTTGTTGA 972
901 TTATGGGCAAGATCACTGGTGTAAATTTCTTCAGAGAAATGGCAAAAGAACTTTGTTGA 960
973 TTATCTAAGAAAGATAGTAAAAAGATATAAG 1005
961 TTATCTAAGAAAGATAGTAAAAAGATATAAG 993

RESULT 3
ACC48532/c
ID ACC48532 standard; DNA; 2981 BP.
XX

AC ACC48532;
DT 11-AUG-2003 (first entry)
XX

DE Staphylococcal surface-exposed immunogenic polypeptide DNA.

XX Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW antibacterial; vaccine; gene; ds.
XX

OS Staphylococcus aureus.

XX Key Location/Qualifiers
FH complement (968. .1951)
FT /*tag= a
FT /product= "SEIP"
XX

PN WO2003020875-A2.

XX 13-MAR-2003.

XX 17-JUN-2002; 2002WO-US019224.

XX 17-JUN-2001; 2001US-0298975P.

XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.

XX Scott DL;

XX WPI; 2003-300870/29.

XX P-PSDB; ABR41865.

XX Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
containing receptors for siderophores or iron-binding ligands, useful for
producing antibodies effective against Staphylococci infection.

XX Claim 4; Page 53; 62pp; English.

XX The present sequence is that of DNA encoding novel Staphylococcus aureus
surface-exposed immunogenic polypeptide (SEIP) D2 SA03. To isolate SEIP
genes, an expression library of *S. aureus* genomic DNA was screened with
anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
amino acid sequences were deduced and analysed for conserved and/or

CC functional domains. The SEIP encoded by the present DNA sequence has
CC sequence homology with the siderophore family of periplasmic binding
CC proteins. The invention provides methods for the identification, cloning,
CC production and recovery of SEIPs. The SEIPs can be used individually, or
CC in combination, to produce anti-staphylococcal antibodies useful in
CC passive or active immunisation strategies to prevent or contain
CC staphylococcal infection. They can also be used to develop diagnostic
CC assays
XX
SQ Sequence 2981 BP; 988 A; 559 C; 386 G; 1048 T; 0 U; 0 Other;
Query Match 56.4%; Score 569; DB 8; Length 2981;
Best Local Similarity 73.3%; Pred. No. 6.8e-95;
Matches 728; Conservative 0; Mismatches 265; Indels 0; Gaps 0;
QY 7 GAATCAGTCGAGAGGTTTAAAAATTTAAAGTGTAAATGGCTTATTGTTGTTTAAATGCA 66
DB 1987 GGAACGATGAGAGGTCCTAAAACTTTTAGTATATTGGGATTAAATAGTTGCCCTTACTTTTA 1928
QY 67 ACTGAGCATGTGGAATAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
DB 1927 GTTCGAGCTTGTGTAATACGGATAATTCAGTAAAAAAGAAATCATCAACTAAAGATACT 1868
QY 127 GTTGAATCAAGCAGCAGAGAGGTACTACGAAGTACCTTAAACACCCCTAAACGCTGTGTT 186
DB 1867 ATTCGGTAAAGATGAAATGGTACAGTAAAGTACCTTAAAGATGCAAAAGCTATCGTT 1808
QY 187 GTTCTGTGATATTCAATTTGTTGATCGTTAGTTAGTGTGTTTAAACCTGTTGGGATA 246
DB 1807 GTATTAGATCTCATTTGCGAGTGCATTTAGCAGCATTTAGACGTTTAAACCACTGTTAT 1748
QY 247 GGGATGATAACAATAAATCGTATTATTAAACCATTAAGAGATAAATTTGGAAAAATAC 306
DB 1747 GCTGATGATGTAAGAAAAAAGCTATCATTTAAACCACTTTAGAGAAAAAATTTGGGGATTAT 1688
QY 307 ACTTCTGTAGGAACAGTAAAGCACTTAACTTAGAAGAACTCAGTAAACCTTAAACACGAT 366
DB 1687 ACTTCTGTAGGTACAGTAAACAGCCAACTTAGAGGAAATAGTAAATTAACCCGGAT 1628
QY 367 TTAATTATTGCTGATATAATAATAGACAAAGGTATTTTAAAGACTTAAATAAAATTCGT 426
DB 1627 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAAATAAGAAATTAACAAATTTGCA 1568
QY 427 CTTAGATGAACTGAAAGTTTCGATGGAGATTATTAATGAATAATTTGATGCTTTTAAA 486
DB 1567 CCAACATATTCATTAAAGAGTTTGTATGGAGACTACAAACAAAAATTAATTCGTTCAA 1508
QY 487 ACAATTTCAAAGCTTTAGGTAAAGAGAGAGGTAAACAAACGCTTAGAAGAACACGAT 546
DB 1507 ACAATTTGCTAAAGCTTTAAATAAGAAAAAAGAGGCAAAACGCTTCTGCGCATGAT 1448
QY 547 AAGAAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAATCAAAAGGTATTGGCT 606
DB 1447 AAATTAATCAATAAGTATAAAGTGAATTAATTTATGATAGAAATCAAAAGTGTCTCA 1388
QY 607 CGAGTAGCTGCTTAATTCAGTTTGTGCTCATCCAGCAACTTTATGTTGGTCAATTC 666
DB 1387 GCAGTAGTTGCTTAAAGCTGGTTTATTAGCACATCCAAACTATTTCATATTTGGACAATTT 1328
QY 667 CTAAGTCAACTAGGTTTAAAGAGCATTAAAGTGATGATGTTTACTAAAGGTTTAAAGTAAG 726
DB 1327 TTAAGCAAACTAGGATTTTAAATAATGCAATTAAGTGACGATGTAACAAAGGTTTAAAGTAA 1268
QY 727 TATCTTAAAGGACCTTACTTCAAAATGAACACTGAACTTTTATCTCAAGTGAATCCTGAG 786
DB 1267 TATTGAAAGGACCTTACTTCAATTTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 1208
QY 787 CTAATGTTTCAATATGACAAACAAAGCAAGTTCTTAAGAACCTTCACTAAAGAACTAGAA 846
DB 1207 CGTATGATCATTTATGACAGATCATGCTTAAAAAAGATTTCTGCTGAAATTCGAAGGTTTACAA 1148
QY 847 AAAGATCCGTGATGGAAGAAATTAACCGCTGTGAAATAATCAACGCTGTTGATATTTTAGAC 906

DB 1147 GAAGATCAACATCGAAAAAGTTGATTCAGTCTAAAAATATATCGGTGATATTGTTGAC 1088
QY 907 CGTGACTTATGGCAAGATCACTCGTGGTTTAAATTTCTTCAGAAAGAAATGGCAAAAGACTT 966
DB 1087 CGTGATGTTTGGCAAGATCTCTGCTGGCTTAATTTCTTCTGAAGAAATGGCTTAAAGAACTT 1028
QY 967 GTTCAATTTATCTAAGAAAGATAGTAAAAAAGAT 999
DB 1027 GTTGAATTTATCAAAAAAAGAAACAAAGTAAGGT 995
RESULT 4
ACA19774
ID ACA19774 standard; DNA; 984 BP.
XX ACA19774;
XX AC
XX 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #1431.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Staphylococcus aureus.
XX PN WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Tamwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
DR P-PSDB; ABU15904.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 7644; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 984 BP; 399 A; 123 C; 181 G; 281 T; 0 U; 0 Other;

Query Match 56.2%; Score 567; DB 8; Length 984;

Best Local Similarity 73.6%; Pred. No. 1.5e-94;

Matches 723; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY	14	TCGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGGTTTAAATGCACTGCAG	73
DB			
QY	2	TCGAGGTTTAAAAATTTAAAGTGAATTTGGCTTATTTGGTTTAAATGCACTGCAG	61
DB			
QY	74	CATGTGGAATTAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGAGTTGAA	133
DB			
QY	62	CTTGTGTAATTCGGTAATTAATCAAGTAAAGAGATCACTCAAAAGATGAGTTGAA	121
DB			
QY	134	TCAAGCAGCAAGAGTACTACGAAAGTACCTAAACACCTAAACGTTGTTGTTCTTG	193
DB			
QY	122	TAAAGATGAAATGTCAGTAAAGTACCTAAAGATGCAAAAGTATCGTTGTTATAG	181
DB			
QY	194	AGTATTCATTTGTTGAGCGTTAGTTGCTTTAGATGTTAACTGTTGGGATGAGCGATG	253
DB			
QY	182	AGTACTCAATTCGAGATCATTAGCAGCATTAACGCTTAAACGAGTTGTTGTTGATG	241
DB			
QY	254	ATAACAAAAATTCGTATTTAATTAACCATTAAGATTAATTCGAAATACACTTCG	313
DB			
QY	242	ATGTTAAGAAAAACGTATCATTAACACCATTAAGATTAATTCGAAATACACTTCG	301
DB			
QY	314	TAGGAACAGTACAGCAACCTCACTTAAAGAAATCAGTAAACCTTAAACAGATTTAATTA	373
DB			
QY	302	TAGGTACAGTAAACAGCCAACTTACAGGAAATTAAGTAAATTAACCCGGATTTAATTA	361
DB			
QY	374	TTGCTGATTAATATAGACACAAAGGTATTTATAAGACCTTAAATAAATTTGCTCTACGA	433
DB			
QY	362	TCGCTGATAGCAGTACAGATTAAGGTATTAATAAGAAATTAACAAATTTGCACCAAT	421
DB			
QY	434	TTGAACGAAAGTTTCGATGAGATTAATGAATTAATTTGATGCTTTTAAACCAATTT	493
DB			
QY	422	TATCATTAAGAGTTTGTGAGGAGCTACAAACAAATATTAATTCGTTCAAAACCAATTTG	481
DB			
QY	494	CAAAAGCTTTAGGTAAGAGAGGTAAGAGGCTTAAAGAGCTTAGAAGAACACGATAAGAAA	553
DB			
QY	482	CTAAAGCTTTAATAAGAAAGAAAGAGGCGAAACGCTTGTGAGCATGATTAATTA	541
DB			
QY	554	TTGAAGATATAAAAAAGAAATCACTATGATTAATAAATCAAAAGGATTTGCTTCGAGTAG	613
DB			
QY	542	TCAATTAAGTAAAGATGAATTAATTTGATAGAAATCAAAAGTGCTTCCAGCAGTAG	601
DB			
QY	614	CTGCTAAATCAGTTTGTCTGCTCATCCAGCAACTCTTATGTTGTCATTTCCCTTAAGTC	673
DB			
QY	602	TTGCTAAAGCTGTTTATTAGCAATCCAAACTTATCATATGTTTGACAAATTTTAAACG	661
DB			
QY	674	AATCAGTTTAAAGAGCAATTAAGTGATGCTTAAAGGTTTAAAGTAAAGTATCTTTA	733
DB			
QY	662	AATCAGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	721
DB			
QY	734	AAGGACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTGAATCTCTGAGCGTATGT	793
DB			
QY	722	AAGGACCTTACTTACAAATTAGACACTGAATTTAGTCTGATTTAAATCCAGAGCGTATGA	781
DB			
QY	794	TCAATATGACAAACAAAGCAAGTTCTTAACGAACCTTCACTAAAGAACTAGAAAAGATC	853
DB			
QY	782	TCATATGACAGATCATGCTTAAAGAAAGATTTCTGCTGAATTCAGAAAGTTACAGAAAGATG	841
DB			

QY	854	CTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTTGATATATTTAGACCGTGACT	913
DB			
QY	842	CAACATGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGATATTTGACCGGTGATG	901
DB			
QY	914	TATGGCAGATCAGTGGTTTAAATTTCTCAGAGAAATGCAAAAGAACTTTGTTGAAT	973
DB			
QY	902	TTTGGCAGATCTCGTGGCTTAAATTTCTTGAAGAAATGCTTAAGAACTTTGTTGAAT	961
DB			
QY	974	TATCTTAAGAAAGATAGTAAAAAA 996	
DB			
QY	962	TATCAAAAAAAGAACAAAAAGTAA 984	
DB			

RESULT 5

AAS54519

ID AAS54519 standard; DNA; 1014 BP.

XX

AC AAS54519;

XX

DT 13-FEB-2002 (first entry)

XX

DE Staphylococcus aureus DNA for cellular proliferation protein #831.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

PN WO200170955-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-US009180.

XX

PR 21-MAR-2000; 2000US-0191078P.

PR

PR 23-MAY-2000; 2000US-0206848P.

PR

PR 26-MAY-2000; 2000US-0207727P.

PR

PR 23-OCT-2000; 2000US-0242578P.

PR

PR 27-NOV-2000; 2000US-0253625P.

PR

PR 22-DEC-2000; 2000US-0257931P.

PR

PR 16-FEB-2001; 2001US-0269308P.

XX

(ELIT-) ELITRA PHARM INC.

XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX

XX WPI; 2001-611495/70.

DR

P-PSDB; AAU36660.

XX

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids.

XX

PS Claim 27; SEQ ID NO 8156; 511pp; English.

XX

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX	SQ	Sequence	1014 BP; 410 A; 126 C; 187 G; 291 T; 0 U; 0 Other;
		Query Match 56.0%; Score 564.4; DB 4; Length 1014; Best Local Similarity 73.1%; Pred. No. 4, 4e-94; Matches 724; Conservative 0; Mismatches 266; Indels 0; Gaps 0	
Qy	7	GAATCAGTGAGAGGTTTAAAAATTTTAAAGTGTAAATGGCTTATTGTTTGTTTTAAATGCA	66
Db	25	GGACGATGAGAGGCTAAAAACCTTTTAGTAGTATTGGGATTAATAGTTGGCTTATTTT	84
Qy	67	ACTGCGATGTGGAAATAATAGTTCAGAGTCACTCAAGTAAAGAGTCAATCAAAAGATGGA	126
Db	85	GTTGCGAGCTTGTGTAAATACGGATAAATCAAGTAAAAAAGAAATCATCAACTAAAGATACT	144
Qy	127	GTTGAAATCAACGACGAGAGAGGTACTACGAAAGTACCTAAACACCCCTAAACGCTGCTGT	186
Db	145	ATTTCGGTAAAGATGAAAAATGGTACAGTAAAGTACCTTAAGATGCAAAACGTATCGTT	204
Qy	187	GTTCTTGAGTATTCAATTTGTTGATGCGTTAGTTGCTTTAGATGTTTAAACCTGTTGGGATA	246
Db	205	GTAATTAGAGTACTCATTTGCGAGATGCTATTAGCAGCATTTAGACGTTTAAACCAAGTTG	264
Qy	247	GCGGATGATAACAAAAAATTCGTATTATTAAACCATTTAAGAGATAAAATTTGGAAAAATAC	306
Db	265	GCTGATGATGGTAAGAAAAACGTCATCATTTAAACCCAGTTAGAGAAAAAATTTGGGAATTTAT	324
Qy	307	ACTTCTGTAGGACACGCTAAGCAACCTAATCTTAGAAGAAATCAGTAAACTTTAAACCCAGAT	366
Db	325	ACTTCTGTAGGTACACGCTAAGCAACGCTAAGGAAATTTAGTAAATTTAAACCCGAT	384
Qy	367	TTAATTTATGCTGATAATTAATAGACACAAAGGTATTTATAAGACTTTAAATAAAAATTCCT	426
Db	385	TTAATTTATGCTGATAGCAGTAGACATTAAGGTATTAATTAAGATTTAAACAAATTTGCA	444
Qy	427	CCTACGATTGAACCTGAAAGTTTCGATCGGAGATTATAATGAATAATTTGATGCTTTTAAA	486
Db	445	CCAAACATTTATCATTTAAAGAGTTTTCGATGGAGACTACAAACCAAAACATTTAAATTCGTTCAAA	504
Qy	487	ACAAATTTCAAAAGCTTTAGGTAAAGACGAAGGTAAACAAACGCTTAGAAGAACACAGAT	546
Db	505	ACAAATGCTTAAAGCTTTTAAATTAAGAAAAAAGAGCGGAAACGCTTTCGCTGAACATGAT	564
Qy	547	AAGAAATTTGAAGATAATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATGCTCT	606
Db	565	AAATTAATCAAAAGCTATTAGATGCAATTAAGTTTGTATGAAATCAAAAGTGTCTTCCA	624
Qy	607	GCAGTAGCTGCTTAAATCAGGTTTGTGCTCATCCAAGCAACTCTTATTTGTTGGTCAATTC	666
Db	625	GCAGTTGTTGCTTAAAGCTGTTTATTAGCACATCCAAACTATTTCATATGTTGGGACAAATTT	684
Qy	667	CTAAGTCAACTAGTGGTTTAAAGAACATTAAGTGATGATGTTACTAAAGGTTTAAAGTAAG	726
Db	685	TTAAACGAACCTTGGATTTTAAAAATGCATTAAGTGATGATGTAAACAAAAAGGTTTAAAGTAAA	744
Qy	727	TATCTTTAAGGACCTTACTTCAAAATGAACACTGAAACCTTTATCTTCAAGTGAATCCTGAG	786
Db	745	TACTTTGAAGGACCTTACTTCAAAATGAACACTTACATCTGAACTTTAGCTGACCTTAAATCCTGAA	804
Qy	787	CGTATGTTTCATAATGACAAAAACAAAGCAAGTTCTTAACGAACCTTTCCTAAAGAACTAGAA	846
Db	805	CGCATGATTAATATGACAGATAATGCTAAAAAAGATTTCTGCTGAAATTCAGAAGTTTACAA	864
Qy	847	AAAGATTCCTGTTATGGAAGAAATTAACCGTGTGAAAAATCAACGTTGTTGATATTTTAGAC	906
Db	865	GAAGATCCAACTTGTGAAAAAGTTGAACCGAGTTAAAAAATAATTCGCGGTGAGATTTGTTGAC	924
Qy	907	CGTGACTTATGGCAAGATCACGCTGTTTAAATTTCTTCAGAGAAATGCAAAAGAACTTT	966
Db	925	CGTGATGTTTGGCGAAGATCTCGTGCTTTAAATTTCTTCTGAAGAAATGGCTTAAGAACTTT	984
Qy	967	GTTGAAATTTATCTAAGAAACATAGTAAAAA	996

D	B		985 GTTGAATTATCAAAAAAGACAAAAGTAA 1014
<hr/>			
R	E	S	RESULT 6
A	C	F	ACF74986
I	D		ACF74986 standard; DNA; 1044 BP.
X	X		AC AC
X	X		ACF74986;
X	X		
D	T		20-NOV-2003 (first entry)
X	X		
D	E		Staphylococcus aureus DNA #2666.
X	X		
K	W		Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
K	W		enzymatic assay; antibiotic target; gene; ds.
X	X		
O	S		Staphylococcus aureus.
X	X		
P	N		WO200294868-A2.
X	X		
P	D		28-NOV-2002.
X	X		
P	F		27-MAR-2002; 2002WO-IB002637.
X	X		
P	R		27-MAR-2001; 2001GB-00007661.
X	X		(CHIR-) CHIRON SPA.
P	A		
X	X		Masignani V, Mora M, Scarselli M;
P	I		
X	X		WPI; 2003-120786/11.
D	R		P-PsDB; ABM73426.
X	X		
P	T		New Staphylococcus aureus protein, useful as a vaccine for treating or
P	T		preventing Staphylococcal infection, specifically an infection caused by
P	T		S. aureus, e.g. sepsis.
X	X		
P	S		Claim 6; SEQ ID NO 5331; 49pp; English.
X	X		
C	C		The invention relates to novel genes and encoded proteins from
C	C		Staphylococcus aureus. A composition comprising the S. aureus protein, a
C	C		nucleic acid encoding the protein, or an antibody to the protein, is
C	C		useful as a pharmaceutical, particularly as a vaccine for treating or
C	C		preventing infection due to Staphylococcus bacteria, specifically an
C	C		infection caused by S. aureus. The composition is particularly useful for
C	C		treating or preventing sepsis in a patient. The composition can also be
C	C		used for diagnostics. The protein is also used in an assay for enzymatic
C	C		studies and as a target for antibiotics. This sequence represents one of
C	C		the novel S. aureus genes of the invention
X	X		
S	Q		Sequence 1044 BP; 425 A; 128 C; 191 G; 300 T; 0 U; 0 Other;
<hr/>			
Q	U		Query Match 55.3%; Score 557.4; DB 8; Length 1044;
			Best Local Similarity 73.5%; Pred. No. 8.5e-93;
			Matches 711; Conservative 0; Mismatches 256; Indels 0; Gaps 0
Q	Y		20 GTTTAAAAATTTTAAGTGTAATTCGCTTATGTGTTTTTAAATTGCCAATCGCAGCATGTG 79
D	b		71 GGCTCAACCCTTTAGTATATATGGATTAATAGTTGCCTTACTTTTAGTCGAGCTTGTG 130
Q	Y		80 GAAATAATAGTTCGAAGTCAAGTAAAGAGTCATCAAAAAGATGGAGTTGAAATCAAGC 139
D	b		131 GTAATACGGATAATTCAGTAAAAAAGAATCATCAACTAAAGATACTATTTCCGGTAAAG 190
Q	Y		140 ACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTTGTTGTTCTTGAGTATT 199
D	b		191 ATGAAAAATGGTACAGTAAAGTACCTAAAGATGCAAAACGATTCGTTGTTATAGAGTACT 250
Q	Y		200 CATTTGTTGATGCGTTAGTTCCTTTAGATGTTAAACCTGTTGGGATACGGATGATAACA 259
D	b		251 CATTTGCAATGCATTACGACCATTAGACGTTAAACCAAGTTGGTATTCGTAATGATGTA 310
Q	Y		260 AAAAAAATCGTATTATTAAACCATTTAAGAGATAAAATTTGGAAAAATACACTTCTGTAGGAA 319

Db 311 AGAAAAACGATATCATTAACACGATTAGAGAAAAAATGGGATTTATCTCTGTAGTA 370
Qy 320 CACGTAAGCAACCTAACTTAGAGAAATCAGTAAATCTTAAACCAAGATTTAAATTTGCTG 379
Db 371 CACGTAAGCAACCTAACTTAGAGAAATTAGTAAATTAACCGGATTTAAATTTGCTG 430
Qy 380 ATAATAATAGACACAAAGGATTTATTAAGACCTTAAATAAAATTTGCTCTAGATTAAC 439
Db 431 ATAGCAGTAGACATAAGGATTTATTAAGAAATTAACCAAAATTTGCCAACATTTATCAT 490
Qy 440 TGAAGATTTTCATGAGATTTATTAAGAAATTTGCTCTTAAACCAATTTCAAAG 499
Db 491 TAAAGATTTTTCATGAGACTACAAACAAATTTAAATTCGTTCAAAACAAATTTGCTAAG 550
Qy 500 CTTTAGGTAAAGAGAGCTTAAACGCTTAAAGCAACCAAGATTAAGAAATTTGAG 559
Db 551 CTTTAAATTAAGAAAGAGCGGAAAAACGCTTCTGCTGAGCATGATTAATTAATCAATA 610
Qy 560 AATATAAAAAAGAAATAACTATGATGATAAAAAATCAAAAGGATTTGCTGCGTAGCTGCTA 619
Db 611 AGTATTAAGATGAATTAATTTGATAGAAATCAAAAGGCTTCCAGCAGTAGTTGCTA 670
Qy 620 AATCAGTTTGTCTCATCAAGCAACTCTTATGTTGGTCAATTCCTTAAGTCAACTAG 679
Db 671 AAGCTGGTTTATTAGCACATCAAACTATTCATATGTTGGCAATTTTAAACGAACTAG 730
Qy 680 GTTTTAAAGAACGATTAAGTGTATGTTACTTAAAGGTTTAAAGTATCTTAAAGGAC 739
Db 731 GATTTTAAATGATTTAAGTGACGATGTAACAAAGGTTTAAAGTAAATTTTGAAGGAC 790
Qy 740 CTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCTGAGCGTATGTTCAATA 799
Db 791 CTTACTTACAAATGAACACTGAACTTTATGTTAGTGAATTTAAATCCAGAGCGTATGATTA 850
Qy 800 TGACAAACAAAGCAAGTTCTTAAGCACTTCACTAAAGAACTTAAAGAAAGATTCCTGTAT 859
Db 851 TGACAGATCATGCTAAAGAAAGATTCCTGCTGAAATTCAGAAAGTTTACAAAGAAAGTCAACAT 910
Qy 860 GGAAGAAATTAACGCTGTGAAATCAACGTTGATATTTTACGCGTATGCTTATGGG 919
Db 911 GGAAGAAATTTGAATGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 970
Qy 920 CAAGATCAGCTGTTTAAATTTCTTCAAGAAATGCAAAAGAACTTTGTTGAATTAATCA 979
Db 971 CAAGATCAGCTGTTTAAATTTCTTCAAGAAATGCTTAAAGAACTTTGTTGAATTAATCA 1030
Qy 980 AGAAGA 986
Db 1031 AAAAGA 1037

RESULT 7

AAS51822

ID AAS51822 standard; DNA; 927 BP.

XX

AC AAS51822;

XX

DT 13-FEB-2002 (first entry)

XX

DE Staphylococcus aureus DNA for cellular proliferation protein #239.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX

KW antibacterial; drug design.

XX

OS Staphylococcus aureus.

XX

XX WO200170955-A2.

XX

XX 27-SEP-2001.

XX

XX 21-MAR-2001; 2001WO-US0009180.

XX

PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haseelbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR P-PSDB; AAU33963.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 4404; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 927 BP; 370 A; 120 C; 169 G; 268 T; 0 U; 0 Other;

Query Match 52.8%; Score 531.8; DB 4; Length 927;

Best Local Similarity 73.4%; Pred. No. 4e-88;

Matches 680; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

Qy 46 TTAATGTTGTTTAAATTCGAACTGCGATCGGAAATTAAGTTCAGTAACCAAGT 105

Db 1 TTAATGTTGCTTATTTTGTGAGCTTGTGTAATACCGATAATTCAGTAACCAAGT 60

Qy 106 AAAGAGTTCATCAAAAGATGGAGTTGAAATCAAGCAGCAAGAGGTAAGTACGAAAGTACCT 165

Db 61 GAATCATCACTAAAGATACATATTTTCGGTAAAGATGAAATGGTACAGTAAAGTACCT 120

Qy 166 AAACACCCCTAAACGTTGTTGTTCTTGAGTATTCATTTGTTGATCGGTTAGTCTTTA 225

Db 121 AAAGATGCAAAACGTTATCGTTATTAAGAGTACTATTTCAGATGCAATTCAGCAGCA 180

Qy 226 GATGTTAAACCTGTTGGATAGCGATGATCAAAATAAAATTCGTTATTAACCAATTA 285

Db 181 GAGCTTAAACCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240

Qy 286 AGAGATAAAATTTGGAAATATACACTTCTGTAGGAAACGCTAAGCAACTCACTTAGAAGAA 345

Db 241 AGAGATAAAATTTGGAAATTTATCTTCTGTAGGTAACGTTAAACGCAACTTAGAGGAA 300

Qy 346 ATCAGTAAATTTAAACCAAGTTTAAATTTGTTGATTAATAATAGACCAAGGATTTAT 405

Db 301 ATTAGTAAATTTAAACCGGATTTAAATTCGTTGATGATGATGATGATGATGATGATGAT 360

Qy 406 AAAGACTTAAATTAATTTGCTCTAGGTTGAAAGTTCGATCGGATTAATTAATTAATTAAT 465

Db 361 AAAGAAATTAACCAAAATTTGCACCAATTTATCATTAAGAGATTTTGTGAGACTACAAA 420

```
QY 466 GAAATATTGATGCTTTTAAACAAATTTCAAAAGCTTTAGCTAAAGAGAGAGAGTAAA 525
DB 421 CAAACACATTAATCGTTCAAAACAAATGCTAAAGCTTTAAATAAAGAAAAAGAGCGGAA 480
QY 526 AAAAGCTTAGAAGAACACGATAGAAAATTTGAAGAATATAAAAAAGAAATAAATATGAT 585
DB 481 AAGCGCTTGTGTAACATGATATAATTAATCAAAAGATATAAGATGAATTAAGTTTGCAT 540
QY 586 AAAAATCAAAAGGATATTCCTGCGAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAAGC 645
DB 541 AGAAATCAAAAGAGTCTTCCAGCAGTTGTTGCTAAAGCTGGTTTATTAGCACATCCAAAC 600
QY 646 AACTCTTATGTTGGTCAATTCCTTAAGTCAACTAGTCACTAGTCTTTAAAGAGCATTAAAGTATGAT 705
DB 601 TATTCATATGTTGGCAATTTTAAACGAACCTTGGATTTAAATAATGCAATTAAGTATGAT 660
QY 706 GTTACTAAAGGTTTAAAGTAAGTATCTTAAAGGACCTTACTTACAATGAACACTGAAACT 765
DB 661 GTAACAAAGGTTTAAGTAAATACCTTGAAGGACCTTACTTACAATTAGACTGAACAT 720
QY 766 TTATCTCAAGTGAATCTGAGCGGTATGTTCTAATGACAAACAAAGAGTTCTTAACGAA 825
DB 721 TTAGCTGACTTAAATCTCGAAGCAGATGATTAATGACAGATAATGCTTAAAAAAGATTCT 780
QY 826 CTTTCACTAAAGAACTAGAAAAGATCCTGTATGGAAGAAATTTAAACGCTGTGAAAT 885
DB 781 GCTGAATTCAAAGAGTTTCAAGAGAGATCCAACTTGGAAAAAGTTGAACGCAAGTTAAAAAT 840
QY 886 CAACGTTGATATTTAGACCGTACCTTATGGGCAAGATCACGTTGTTTAAATTTCTTCA 945
DB 841 AATCGCTGGATATTTGACCGTGATGTTGGGCAAGATCTCGTGGCTTAATTTCTTCT 900
QY 946 GAAGAAATGGCAAAAGAACTTTGTTGAA 972
DB 901 GAAGAAATGGCTTAAAGAACTTTGTTGAA 927
```

RESULT 8

```
AAT80398/c
ID AAT80398 standard; cDNA; 2247 BP.
XX AC AAT80398;
XX DT 14-APR-1998 (first entry)
XX DE Staphylococcus aureus Gene #4 encoding cDNA sequence 1.
XX KW Staphylococcus aureus WCUH 29; antagonist; antibacterial; immunogen;
XX KW vaccine; disease; protection; isolation; ss.
XX OS Staphylococcus aureus.
XX PN WO9731114-A2.
XX PD 28-AUG-1997.
XX PF 25-FEB-1997; 97WO-GB000524.
XX PR 26-FEB-1996; 96GB-00004045.
XX PA (SMK ) SMITHKLINE BEECHAM PLC.
XX PI Burnham MKS, Hodgson JE;
XX PS WPI; 1997-435166/40.
XX DR New Staphylococcus aureus polynucleotide and polypeptide(s) - for
XX PT isolating antagonist of the polypeptide(s) useful as anti-bacterials.
XX PS Claim 4; Page 30-31; 117pp; English.
XX CC The present sequence encodes a novel polypeptide, which is optionally
```

expressed in NCIMB 40771. The polypeptide, and polynucleotide encoding it, are derived from *Staphylococcus aureus*. Cells expressing ligands binding the polypeptide can be used to isolated candidate compounds that bind and inhibit the activity of the polypeptide. Such compounds can be used as anti-bacterial compounds. The polypeptide may also be used as an immunogen to vaccinate an animal for protection against *Staphylococcus aureus* caused disease

Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;

Query Match 33.4%; Score 336.2; DB 2; Length 2247;
Best Local Similarity 70.6%; Pred. No. 2.5e-52;
Matches 580; Conservative 0; Mismatches 229; Indels 13; Gaps 10;

```
QY 1 GGATGGGAATCAGTGAGAGGTTTAAAAATTTTAAAGTGAATTTGGCTTATTTGTTTGA 60
DB 831 GGAGTGGGACNGATGAGAGGTTCTAAAAACCTTTTAGTATATTGGGATTAATAGTTGCC 772
QY 61 ATTGCAACTGCAGCATGTGGAATA--ATAGTTCAAGTAACTCAAGTAAAGAGTCATCA 117
DB 771 CTTTATGTTGAGCTTTGGTAAATACGGATAATTCAGTAAAGAG-AATCATCAACTA 713
QY 118 AAGATGAGGTGAAATCAAGCAGAAAGAGGTACTAGAAAGTACCTAAACACCCCTAAA 177
DB 712 AAGATACTATTTCGGGTAAAAAGATCAAAATCGTACAGTAAAAAGTACCTTAAAGATGCA 653
QY 178 CGTGTGTTGTTCTTGTAGTATTCATTTG--TTGATCGGTTAGTTGC-TTTTATAGTTTAAA 234
DB 652 CGTATCGTTGTTATTAGAGTACTCATTTTGCAGAGTTTGCATTANCAGCATATAGCCGTT 593
QY 235 CC-TGTTGGGATAGCGGATGATAACAAAAAAATCGTATTATTATTAACCATTAAGAGATAA 293
DB 592 CCAAGTTGGTATTGCTGATGATGGTAAAGAAAAACGTTATCA-TAACCAGTAAGAGA-AA 535
QY 294 AATTGGAAATACACTTCTGTAGGAACACGTAAGCAACTCACTTCTAGAGAAATCAGTAA 353
DB 534 AATTGGGGGTAACTCTTCTGTAGGTACACGTAA-CAAGCNAACCTTAAGAGANATTAGTAA 476
QY 354 ACTTAAACCATTTAATTATTGCTGATAATAATAGACACAAAGGTATTATTAAAGACTT 413
DB 475 ATTAANCCGGGATTAATTATCNCCTGATAGCAGTAGACATAAGGGTATTAAATAAGAAT 416
QY 414 AATAAAATTCCTCTACGATGAACTGAAAGTTTCGATGGAGATTATAATGAAATAT 473
DB 415 AGCCCACTATGACCC-ACATTTATCATTAAAGAGTTTGTATGGAGACTACACCCCAATAT 357
QY 474 TGATGCTTTTAA-AACAATTTCAAAAGCTTTAGGTAAAGAGAAAGAAAGCGTAAAAACGCT 532
DB 356 TAAATTCGTTCAACAACTTCTAAAGCTTTAAATATAAGANAAGAAAGCGGNNAAACGTC 297
QY 533 TAGAAGAACACGATAAGAAAATTGAGNATATAAAAAAGAAATTAATGATGATAAAAATC 592
DB 296 TTGCTGAGCATGATAAAATTAATCAATAAGTATAAGATGAATTAATTTGATAGAAATC 237
QY 593 AAAAGGTATTGCTGCGAGTAGCTGATAATCAGGTTTCTGCTCATCCAGCAACTCTT 652
DB 236 ACAAGTCTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCACACTATTCTAT 177
QY 653 ATGTTGGTCAATTCTTAAGTCAACTAGGTTTAAAGAGAGCAATTAAGTATGATGTTACTA 712
DB 176 ATGTTGGCAATTTTAAACGANCCTAGGATTTAAAAATGCAATTAAGTGACGATGAACAC 117
QY 713 AAGGTTTAAAGTATCTTAAAGGACCTTACTTACAATGACACTGAAAGCTTTTATCTC 772
DB 116 AAGGTTTAAAGTAAATTAATTTGAAGGACCTTACTTACAATTAAGTATGACATGCTA 57
QY 773 AAGTGAATCTCGAGCGTATGTTTCAATTAATGACAAACAAAGCAA 814
DB 56 ATTTAAATCCAGAGCGTATGATCATTAATGACAGATCATGCTA 15
```

RESULT 9
AAT83786/c

ID AAT83786 standard; DNA; 2247 BP.
 AC AAT83786;
 XX 16-JUL-1998 (first entry)
 XX DNA encoding 3 Staphylococcus aureus proteins of unknown function.
 DE Staphylococcus aureus protein; ribozyme; antisense sequence; control;
 KW Staphylococcus aureus; regulatory element; bacterial gene expression;
 KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
 KW toxic shock syndrome; ss.
 XX Staphylococcus aureus.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 CDS complement (1212..1445)
 FT /*tag= a
 FT CDS complement (1456..1587)
 FT /*tag= b
 FT CDS complement (1705..2016)
 FT /*tag= c
 PT
 XX WO9730070-A1
 XX 21-AUG-1997.
 PF 19-FEB-1997; 97WO-US002318.
 XX
 PR 20-FEB-1996; 96US-0011888P.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;
 PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
 XX WPI; 1997-424969/39.
 DR P-PSDB; AAW27819, AAW27820, AAW27821.
 XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used to
 PT isolate antimicrobial compounds, and in vaccines against S. aureus
 PT infection.
 XX Claim 9; Page 672-673; 989pp; English.
 XX
 CC The present sequence encodes 3 Staphylococcus aureus proteins of unknown
 CC function. The present sequence was isolated from a library of clones of
 CC S. aureus WCUH 29 in Escherichia coli. The DNA sequence can be used in
 CC the construction of ribozymes and antisense sequences to control the
 CC expression of Staphylococcal genes. The DNA sequence is also useful as a
 CC source of regulatory elements for the control of bacterial gene
 CC expression. The encoded protein may be used to produce vaccines to enable
 CC a host to produce specific antibodies with antibacterial action. These
 CC vaccines and antibodies would protect a host against invasion by S.
 CC aureus, and conditions relating to Staphylococcal infection, e.g.
 CC Staphylococcal food poisoning, scaled skin syndrome, and toxic shock
 CC syndrome
 XX
 SQ Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;
 Query Match 33.4%; Score 336.2; DB 2; Length 2247;
 Best Local Similarity 70.6%; Pred. No. 2.5e-52;
 Matches 580; Conservative 0; Mismatches 229; Indels 13; Gaps 10;
 QY 1 GGAGTGGAAATCAGTGAGAGGCTTTAAATAATTAAGTGCTAAATGGCTTTATTGTTTGA 60
 DB 831 GGAGTGGGACNGATGAGAGGCTCTAAATACTTTTAGTATATTTGGATTAAATAGTTGCCTTA 772
 QY 61 ATTGCAACTGAGCATGTGGAATAAATAATGTTCAAGTAAGTCAAGTAAGAGTCAATCA 117
 DB 771 CTTTATTGAGCTTGTGTAATACCGGATAATTCAGTAAAGAAAGAG-AATCATCAACTA 713
 QY 118 AAGATGGAGTTGNAATCAAGCAGGAGAGGTTACTACGAAAGTACCTAAACACCCCTAAA 177

Db 712 AAGTACTATTTCGGGTAAAGATGAAATGGTACAGTAAAGTACCTAAAGATGCACAA 653
 QY 178 CGTGTGTTGTTCTTGGAGTATTCAATTTG- -TTGATGCGTTAGTTGTC- -TTTAGATGTTAAA 234
 Db 652 CGTATCGTTGTTAGAGTACTCATTTGCAAGTTGCAATTANCAGCATATAGCCGTTANA 593
 QY 235 CC-TGTTGGGATAGCGGATGATAACAAAAAATCGTATTATTAAACCATTTAAGAGATAA 293
 Db 592 CCAAGTTGGTATTGCTGATGATGGTAAAGAAAAACGTATCA- -TAACCCAGTAAGAGA-AA 535
 QY 294 AATTGGAAAAATACACTTCTGTAGGACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
 Db 534 AATTGGGGGTAAATCTTCTGTAGGTACACGTAA- -CAAGCNAACCTTAAGAGANAATTAGTAA 476
 QY 354 ACTTAAACACGATTTAAATTATTGCTGATAATAATAGACACACAAAGGTATTTATAAGACTT 413
 Db 475 ATTAAACCGGATTAATTATCNCGTAGTAGCAGTAGACATAAGGGTATTATTAAGAAATT 416
 QY 414 AATAAAAAATGCTCTACGATTGAACCTGAAAGTTTCGATGGAGATTATATGAAATAAT 473
 Db 415 AGCCACATTTGCACC-ACATTATCATTTAAAGAGTTTTCATGGAGACTACACCCACAATAT 357
 QY 474 TGATGCTTTTAA- -AACAAATTTCAAAGCTTTTAGGTAAGAGAAAGAAAGCTAAACACGCT 532
 Db 356 TAAATTCGTTCAACAAATTTGCTAAAGCTTTTAAATTAAGANAAGAAAGCGNNAACGTC 297
 QY 533 TAGAAGAACACGATAGAAATAATTGAAGAAATATAAAAAAGAAATAACTATGGATAAAAAATC 592
 Db 296 TTGCTGAGCATGATAAATAATCAATAAGTATTAAGATGAATTAATTAATTTGATAGAAATC 237
 QY 593 AAAAGGTATTGCTGCTGAGTAGCTGCTAAATCAAGTTTGTCTCTCATCAAGCAACTCTT 652
 Db 236 ACAAGTGTCTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCACACTATTCTAT 177
 QY 653 ATGTTGGTCAATTCCTAAGTCAACTAGTAGTTTAAAGAGCACTTAAGTGTATGTTACTA 712
 Db 176 ATGTTGGACAATTTTAAACCGANCTAGGATTTTAAAAAATGSCATTAAGTGCAGATGAACAC 117
 QY 713 AAGGTTTAAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTC 772
 Db 116 AAGGTTTAAAGTAATAATTTTGAAGGACCTTACTTACATTAAGACACTGAACATTAGCTG 57
 QY 773 AAGTGAATCCTGAGCGTATGTTTCATAATGACAAAAACAAACAA 814
 Db 56 ATTTAAATCCAGCGTATGATCATTTATGACAGATCATGCTA 15
 RESULT 10
 ID ACA47237
 XX ACA47237 standard; DNA; 525 BP.
 AC ACA47237;
 XX 19-JUN-2003 (first entry)
 DT Prokaryotic essential gene #28894.
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX Staphylococcus haemolyticus.
 OS WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.

CC	assays	
XX	Sequence 2957 BP; 1160 A; 381 C; 512 G; 904 T; 0 U; 0 Other;	
SQ	Query Match 26.7%; Score 269.2; DB 8; Length 2957;	
	Best Local Similarity 55.4%; Pred. No. 4.6e-40;	
	Matches 545; Conservative 0; Mismatches 433; Indels 6; Gaps 1;	
QY	17 GAGGTTTAAATTTTAAAGTAAATGGCTTATGTTGTTTAAATGCAATGCGAGCAT 76	
Db	990 GAGGTTTAAATTTTAAAGTAAATGGCTTATGTTGTTTAAATGCAATGCGAGCAT 1049	
QY	77 GTGGAAATATAGTTCAGTAACTCAAGTAAAGAGTCAATCAAAAGATGCGATTGAAATCA 136	
Db	1050 TAGCAGTACGGGTTGTGTCACAAAGATCTGAGAGAAACTGAAATGACGACATAA 1109	
QY	137 AGCAGCAAGAGTACTACGAAAGTACCTAAACACCTTAAACGTTGTTGTTCTTGAAT 196	
Db	1110 AAGATGAAATAGCACTGAAAGAAATTAAGAAAGAAATCTTAAACGTTGTTGTTGTTAGAAAT 1169	
QY	197 ATTCATTTGTTGATGCTTGTAGTTGCTTAAAGCTGTTAAACCTGTTGGGATAGCGGATGATA 256	
Db	1170 ATAGTTTTCGCTGATATTTAGCAGCAATAGATGAAACCTGTTGGTATTTGCAGATGATG 1229	
QY	257 ACAAAAAAATCGTATTATTAAACCAATTAGAGATAAAATTTGGAAATATACACTTCTGTAG 316	
Db	1230 GCAGCACTTAAATATAACAAAGTCACTAAGAGATAAGATTTGGGCATATGAATCGTTG 1289	
QY	317 GAACAGTAAAGCAACTACTTAGAGAAATCAGTAAACCTTAAACAGATTTAAATTTATG 376	
Db	1290 GATCTAGACCGCAACCGAATATGGAAGTGAATGAATTAATAACCGGATTTGATCATTTG 1349	
QY	377 CTGATTAATATAGACACAAAGGTATTATTAAGACCTTAAATTAATTTGCTCTACGATTG 436	
Db	1350 CAGATTTAGCAGACATAGAAATCAAAATCAGAAATTTGAGCAAAATTTGCTCCGACAAATCA 1409	
QY	437 AACTGAAAGTTTCGATGAGATTATAATGAAATATTCATGCTTTTAAACAAATTTCAA 496	
Db	1410 TGTATTAGCGGTACGGGAGATTATTAATGCAATTTTGAAGCATTTTAAACAGTCGCTA 1469	
QY	497 AAGCTTTAGGTAAAGAGAGAGGTAAAGAAACCGCTTAGAAGAACAGATAGAAATTTG 556	
Db	1470 AAGCAGTAGCAAGAGAGAGAGCGGAGAGCGTCTGGAAGAGCATGATAAAATATTAG 1529	
QY	557 AAGAAATATAAAGAGAAATTAATATGATTAATAATCAAAAGTATTGCTCGCAGTAGCTG 616	
Db	1530 CGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAATCTGCAATTTGCAATTCGTTATCT 1589	
QY	617 CTAATCAGCTTTGCTTGCTCATCCAGCAACTCTTATGTTGTCATTCATTCCTAAGTCAAC 676	
Db	1590 CAAGACAGGTATGTTTATTAATTAATGAAGATACATTTATGGACAAATCTTAAATTAATA 1649	
QY	677 TAGGTTTAAAGAGCATTAAAGTGATGTTTACTAAAGGTTTAAAGTAAATGATCTTAAAG 736	
Db	1650 TGGGTATTCAACTGAAGTCACAAAAGACAAAACTACGATGTTGGTGAACGCAAGGGTG 1709	
QY	737 GACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTGAATCTCGACGCTATCTTCA 796	
Db	1710 GTCCCTTATATATTTTAAATTAATGAAAGAACTTGCCCAATATCAATCCAAAGTTGATTT 1769	
QY	797 TAATGACAAACAAAGCAAGTTCTTAAACGAACTTCACTAAAGAACTAGAAAAAGATCCCTG 856	
Db	1770 TAGCCACTGACGAAAGAAACGACAAATA-----TAGACGAATTCATTTGATCTCTGCAG 1823	
QY	857 TATGGAAGAAATTAACGCTGTGAAAGAAATCAACGCTGTTGATATTTTATAGCCGCTGACTTAT 916	
Db	1824 TTTGGAAATCATTTAAAGCTGTGAAAGATTAACAAAGTTTATGACGTTGACCGAAATAAGT 1883	
QY	917 GGGCAAGATCACGTTGTTTAAATTTCTTCAGAGAAATGCGAAAGAACTTTGTTGAATTTAT 976	
Db	1884 GGTGTAATCAAGGGGATTTATCGCAAGTGAAGTATGGCAGAAAGATTTAGAAAAAATTTG 1943	
QY	977 CTAAGAAAGATAGTAAAAAAGATA 1000	
Db	1944 CAGAAAAAGCAAAATAAAAAATACA 1967	
DE	Staphylococcus aureus DNA for cellular proliferation protein #699.	
KW	Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;	
KW	antibacterial; drug design.	
OS	Staphylococcus aureus.	
PN	WO200170955-A2.	
PD	27-SEP-2001.	
PF	21-MAR-2001; 2001WO-US009180.	
PR	21-MAR-2000; 2000US-0191078P.	
PR	23-MAY-2000; 2000US-0206848P.	
PR	26-MAY-2000; 2000US-0207272P.	
PR	23-OCT-2000; 2000US-0242578P.	
PR	27-NOV-2000; 2000US-0253625P.	
PR	22-DEC-2000; 2000US-0257931P.	
PR	16-FEB-2001; 2001US-0269308P.	
XX	(ELIT-) ELITRA PHARM INC.	
XX	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;	
PI	Yamamoto RT, Xu HH;	
XX	WPI: 2001-611495/70.	
DR	P-PSDB; AU36528.	
XX	New polynucleotides for the identification and development of	
PT	antibiotics, comprise sequences of antisense nucleic acids.	
XX	Claim 27; SEQ ID NO 8024; 511pp; English.	
PS	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the genes,	
CC	their use in the discovery of novel antibiotics, the essential genes	
CC	themselves and the encoded proteins. The prokaryotes used are Escherichia	
CC	coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,	
CC	Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also	
CC	useful for the identification of potential new targets for antibiotic	
CC	development. The antisense nucleic acids can also be used to identify	
CC	proteins used in proliferation, to express these proteins, and to obtain	
CC	antibodies capable of binding to the expressed proteins. The proteins can	
CC	be used to screen compounds in rational drug discovery programmes. The	
CC	antisense nucleic acid sequence is also useful to screen for homologous	
CC	nucleic acids which are required for cell proliferation in a wide variety	
CC	of organisms. The present sequence encodes an essential prokaryotic	
CC	cellular proliferation protein. Note: The sequence data for this patent	
CC	did not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 960 BP; 385 A; 115 C; 209 G; 251 T; 0 U; 0 Other;	
SQ	Query Match 26.3%; Score 265.6; DB 4; Length 960;	
	Best Local Similarity 57.8%; Pred. No. 1.9e-39;	
	Matches 494; Conservative 0; Mismatches 354; Indels 6; Gaps 1;	
QY	132 AATCAAGCAGCAAGAGGTACTACGAAAGTACTTAACACCTTAACCGTTGTTGTTCT 191	


```
Db 105 AATAAAGATGATTAAGGAACCTGAAATAATTAAGAAAAATCTAAACGCTGTTGTTGTTAT 164
Qy 192 TGAGTATTCATTTGTTGATGCGTTAGTTCGTTTGTAGATGTTAAACCTGTTGGGATAGCGGA 251
Db 165 AGAATATAGTTTTCGTCGTTATTAGCAGCATTAGATATGAAACCTGTTGGTATTCGAGA 224
Qy 252 TGATAACAAAAAATCGTTATTATTAACCATTAAGAGATATAAATGGAAATACACTTC 311
Db 225 TGATGGCAGCTGATAAATAATTAACAAGTCTAGTAAGATAGATAGGTTGGGCAATTAATC 284
Qy 312 TGTAAGAACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAACTTTAAACACGATTTAAT 371
Db 285 GGTGGATCTAGACCGCAACCGAATATGGAAGTATAGTAATTAATAAACCGATTTGAT 344
Qy 372 TATTCCTGATATAATAGACACAAAGGTTATTTATAAGACTTTAAATAAAATTCCTCTAC 431
Db 345 CATAGCAGATGTTAGCAGACATAAGAAAAATCAATCAGAAATTAAGCAAAATTCCTCCGAC 404
Qy 432 GATTGAACGTAAGATGTTGATGGAGATTTATTAATGAATAATATGATGCTTTTAAACAAT 491
Db 405 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAATATTTGATGCAATTTAAACAAGT 464
Qy 492 TTCAAAAGCTTTAGGTAAAGAAAGAGAGGTTAAAAACGCTTAGAAGAACACGATAAGAA 551
Db 465 CGCTAAGCAGTAGGCAAGAGAAAGAGAGGCGAGAAACGCTCGGAAGACGATGATAAT 524
Qy 552 AATTGAAGATATAAAAAAGAAATTAACATATGATGATAAAATCAAAAGGTTATTCCTCGAGT 611
Db 525 ATTACGGAGATTAGAAAGAAATTTGAACAGAGTACGTTAAAAAATCTGCATTTGCATTCGG 584
Qy 612 AGCTGCTAATCAGTTTTCGTCATCCCAAGCACTTTATGTTGTCATTCCTTAAG 671
Db 585 TATCTCAAGAGCAGGTATGTTTATTAAATGAAGATACATTTATGGGACAAATCTTAAAT 644
Qy 672 TCAACTAGTGTTTAAAGAGCAATTAAGTCATGATGTTACTAAAGGTTTAAAGTAAGTATCT 731
Db 645 TAAATGGGTATTCACCTGAAGTCACAAAGACAAAACTAGCGCATGTTGGTGAACGCA 704
Qy 732 TAAAGACCTTACTTCAAAATGAACACTGAAACTTTTATCTCAAGTGAATTCCTGAGCGTAT 791
Db 705 GGGTGGTCTTATTTATTTAAATGAAGAACTTGCAATATCAATCCAAAAGTTAT 764
Qy 792 GTTCATATGACAAACAGCAAGTTCATACGAACCTTCACATAAGAACTAGAAAAGA 851
Db 765 GATTTAGCCACTGACGGAACCGACAAATAAGTATG-----AACGAAATTCATTTGATCC 818
Qy 852 TCCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTTGATATTTTAGACCGTGA 911
Db 819 TCCAGTTTGGAAATCATTTAAAGCTGTGAAAGATACAAAGTTTATGACGTTGACCGAAA 878
Qy 912 CTTATGGGCAAGATCAGCTGGTTTAAATTTCTTCAAGAAATATGGCAAGAACTTTGTTGA 971
Db 879 TAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTCAAAAGTATGCGGCAAGATTTAGAAAA 938
Qy 972 ATTATCTAAGAAAG 985
Db 939 AATTGCAGAAAAAG 952
```

RESULT 13

AAS51612

ID AAS51612 standard; DNA; 957 BP.

XX

AC AAS51612;

XX

DT 13-FEB-2002 (first entry)

XX

DE Staphylococcus aureus DNA for cellular proliferation protein #29.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX

KW antibacterial; drug design.

XX

OS Staphylococcus aureus.

```
XX WO200170955-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0289308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR P-PSDB; AAU33753.
XX
XX New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Claim 27; SEQ ID NO 4194; Slipp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence encodes an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 957 BP; 382 A; 114 C; 209 G; 252 T; 0 U; 0 Other;
```

```
Query Match 26.2%; Score 264.6; DB 4; Length 957;
Best Local Similarity 57.8%; Pred. No. 3e-39;
Matches 493; Conservative 0; Mismatches 354; Indels 6; Gaps 1;
Qy 132 AATCAAGCACCAAGGAGTACTACGAAAGTACCTAAACACCTAAACCTGTTGTTGTTCT 191
Db 111 AATAAAGATGAATTAGGAAGTGAATAATTAAGAAATCTTAACCTGTTGTTGTTAT 170
Qy 192 TGAGTATTCATTTGTTGATGCGTTAGTTCGTTTATAGATGTTAAACCTGTTGGATAGCGGA 251
Db 171 AGAATATAGTTTTCGTCGATTATTTAGCAGCATTAGATATGAACCTGTTGGTATTCAGA 230
Qy 252 TGATAACAAAAAATCGTTATTTTAAACCATTAAGAGATATAATTTGGAATACACTTC 311
Db 231 TGATGGCAGCAGTAAAAATATAACAAGTCTAGTAAGATAGAGTGGGTCATATGAATC 290
Qy 312 TGTAAGACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAACTTAAACACAGATTTAAT 371
Db 291 GGTGGATCTAGACCGCAACCGAATATGGAAGTATGTAAGTAATTAATAAACCGATTTGAT 350
Qy 372 TATTGCTGATATAATAGACACAAAGGTTATTTATAAGACTTTAAATAAAATTCCTCTAC 431
Db 351 CATAGCAGATGTTAGCAGACATAAGAAAAATCAAAATCAGAAATTAAGCAAAATTCCTCCGAC 410
```


Db 915 AAGTATGCCACAAGATTAGAAAAAATTGCAGAAAAAG 952

RESULT 16

ACA20079

ID ACA20079 standard; DNA; 960 BP.

1000

ACA20079;
19-JUN-2003 (first entry)

DE Prokaryotic essential gene #1736.

XX Antisense; ds; prokaryotic essential gene; cell proliferation; prokaryotic essential gene #1750.
KW Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.

[illegible]

Staphylococcus

XXXXXXXXXX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

21-MAR-2002; 2002WO-US009107.

21-MAR-2001: 2001US-00815242

PK ZI-PBK-2001; Z001US-00948
PR 06-SEP-2001: 2001US-00948

PR 25-OCT-2001: 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-

XX

PA	(ELIT-) ELITRA PHARM INC.		
XX			
XX	Wang L, Zamudio C,	Malone C,	Hasselbeck R, Ohlssen KL, Zyskind JW;
PI	Wall D,	Trawick JD,	Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX			
XX	WPI; 2003-029926/02.		
DR	P-PSDB; ABU16209.		
DR			
XX			
XX			
PT	New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to		
PT	isolate candidate molecules for rational drug discovery programs.		
PT			

PS Claim 14: SEO ID NO 7949: 1766pp: English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 897 BP; 293 A; 176 C; 195 G; 233 T; 0 U; 0 Other;

Query Match 15.5%; Score 156.4; DB 8; Length 897;
Best Local Similarity 51.9%; Pred. No. 1.8e-19;
Matches 454; Conservative 0; Mismatches 381; Indels 39; Gaps 3;

QY 123 TGGAGTTGAATCAAGCAGCAAGAGGTACTACGAAAGTACCTAAACACCTTAACCGTGT 182
DB 63 TCGTGTACCGTCAAGATCAAAAAGGGAATTTACCTTAGATAGCGTACCTAAACGTTG 122
QY 183 TGTGTGTTCTTGAGTATTCATTGTTGATCGGTTAGTTGCTTTAGATGTTAAACCTGTTGG 242
DB 123 TGTGCTTAGAATATTCATTGTTGATGACCTTGACAAATGTTGTTCAGCCCTGTTGG 182
QY 243 GATAGCGGATGATAACAAAAAATCGTATTATTAAACCATTAAGAGATATAAATTGGAAA 302
DB 183 TGTCCGAGATGACATGATAAAACCGGTATTCGCAAAAAGTACGCGATAAAAGTGCAGCC 242
QY 303 ATACACTTCTGTAGGAACAGTAAAGCACTTAAGCACTTAAGTAAAGTAACTAGTAACTTAAC 362
DB 243 ATGGGAATCAGTGGGACGCTTCTCAACCGAGTTTGAAGCGATTTCTGCACTTAAACC 302
QY 363 AGATTAAATATTGCTGATAAATAAGACACAAAGGTATTATAAGACCTTAAATAAAT 422
DB 303 AGATTGATCAITGCCGATGATANTGCCATTCTGCCGCTCTATGAAGAACTCAAAAAT 362
QY 423 TGCTCTACGATTTGAATGAAAGTTTCGATGAGATTTAATGAAATATTGATGCTTT 482
DB 363 CGCGCCGACAGTCGTCTTAAATTCGCCCATGAGAACTATCAAGAAACCTTGAAACCGC 422
QY 483 TAAACAAATTTCAAGCTTTAGGTAAAGAAAGAGAGAGTAAAGAAACCGTTAGAAACA 542
DB 423 ACAAATAATCGGTGATTTATTAGTAAATCAAAAGAAATGCAAGCGCTATTGCAAAACA 482
QY 543 CGATAGAAATTTGAAGATATAAAAGAAATTAATGATGATAAAATCAAAAGTATT 602
DB 483 TAAGCAGGATTTGCGG-----ACATCGCAAAACGTTACCGAAAGGGAAGAAACGAT 536
QY 603 GCCTGACGATGCTGCTTAAATCAGGTTTGCTTGCATCCAAAGCAACTCTTATGTTGGTCA 662
DB 537 TATCGGTGTTTACGTGAAACCCAAATTAATTTATATATAGGAACTCTATGCTGTGG 596
QY 663 ATTCCTAAGTCAATAGGTTTAAAGACGATTAAGTATGATGTTACTAAAGGTTTAA 722
DB 597 CTTAGTGAAGTGCTAGGTTTATCAATGCAAAAGCCGTCGCGATAAAGCAACTTAATGC 656
QY 723 TAAGTATCTTAAGGACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTGAATCC 782
DB 657 TTCGGTTGGTTTGAACAAG-----TGGCGCAAAAGGCC 692
QY 783 TGAGCGTATGTTTCATATGACAAACAAAGCAAGTCTTCAAGCACTTCACTAAAGAACT 842
DB 693 TGATCTGATGATCTTAATCCATTA-----TCGTGATGAGAGTATTGCAAGAAATG 743
QY 843 AGAAAAAGATCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGGTTGATATTTT 902
DB 744 GGAATAATGAAGCCTTATGGAATAATTATCCCTCGCGGTAAATAAATGGTCAATCTTAC 803
QY 903 AGACCGTACTTATGGCAGATCAGTGGTTTAAATTTCTTCAGAGAAATGCAAAAGA 962
DB 804 TAAATGATTAATTTATGGCAGAGACAGTGGTATTTGATGCGCGCTGAAGTAAATGCTAA 863
QY 963 ACTTGTGTAATTTATCAAGAAAGATAGTAAAAA 996

DB 864 AGTCCAAAGACTTTGTAAACGAAATCCGCCAAATAA 897
RESULT 22
ABK74775
ID ABK74775 standard; DNA; 801 BP.
XX
XX ABK74775;
DT 13-AUG-2002 (first entry)
XX
XX Bacillus licheniformis genomic sequence tag (GST) #2066.
DE Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
XX Bacillus licheniformis.
XX WO200229113-A2.
PN 11-APR-2002.
PD
XX 05-OCT-2001; 2001WO-US031437.
PF
XX 06-OCT-2000; 2000US-00680598.
PR 27-MAR-2001; 2001US-0279526P.
XX (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX Berka R, Clausen IG;
PI WPI; 2002-416684/44.
XX
XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
XX Claim 4; SEQ ID NO 2066; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 801 BP; 257 A; 166 C; 204 G; 174 T; 0 U; 0 Other;
Query Match 14.8%; Score 149.4; DB 6; Length 801;
Best Local Similarity 51.8%; Pred. No. 3.4e-18;
Matches 414; Conservative 0; Mismatches 376; Indels 9; Gaps 3;
QY 176 AACGTTGTTGTTGTTGATGATTTCAATTTGTTGATCGGTTAGTTGCTTAGTTAAAC 235
DB 1 AAAAAGTTGTTGTTGTTGATTTTGAATTTGGGATTTTATTTGATCGCTCCTTGTATCGGGTATTAAAGC 60

Qy 236 CTGTTGGATAGCGGATGATCAACAAAAAATCGTATTATTAACCATTAAGAGATAAAA 295
Db 61 CTGTCGGAATTCGCGACGCGCAACCTAAGTTTATTAACGAGAAGGTCTAGGGGAAAA 120
Qy 296 TTGGAAATACACTCTCTAGGACACGTAAAGCAACCTAATCTTAGAAGAAATCAGTAAAC 355
Db 121 TCAAAGGTATACITTCAGTCGGTTCGCGCCGACCAAGCTTTGAAAAAATTCGTTCTT 180
Qy 356 TTAACACAGATTTAATTTCTGATATAATAGACACAAAGGTATTTATAAGACTTAA 415
Db 181 TAAGCCGATTAATTTATTCGGATTTCGACGACGACACGCGCTCTATGTAAGCTGT 240
Qy 416 ATAAATTCCTCTACGATGAACTGAAAAATTTTCGATGGAGATTAATGAATAATGT 475
Db 241 CGAAATTCGCGACCAATCGCGTCAAAAACTTGAATGCGGCTATTCAGGACGCGTTG 300
Qy 476 ATGCTTTTAAACAAATTTCAAGCTTTAGGTAAAGAGAGAGAGGTAAGAAACGCTTAG 535
Db 301 ATGCATCTCTTACGATTTCGAAGCGCTCGGCAAGAAAGCGCAATGGAGAAAAATTTGG 360
Qy 536 AAGAACGATGAAGAAAAATGAAGATATAAAAAAGAAATAACTATGATAAAAATCAAA 595
Db 361 CTGAACATAACAAAAGCTGGATGAATTTGAACAGAAATTCGGACGCGGAAACAGAGCA 420
Qy 596 AGGTATTCGCTGAGTAGCTGTAATCAGGTTTGTCTGCTCATCAAGGCAACTCTTATG 655
Db 421 TTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTCGCGGATGAA---AACTTTTCA 477
Qy 656 TTGGTCAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTAAGTATGATGTTACTAAAG 715
Db 478 CGTCTCAGCTTCTGACGAAATTCGGCTATACATACGGTGTGCGGACAGCG---GCAAG 534
Qy 716 GTTTAAGTAAGTATCTTAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAG 775
Db 535 GCGATGCGGAAAACGGTGAATCCGTCATATTAATAATGACGCTCGAACAGCTGTTGAGA 594
Qy 776 TGAATCTGAGCGTATGTTCTAATATGAC---AAACAAGAGAGTTCTTAACGAACTTCAC 832
Db 595 AAGATCCGAGCTTATTCCTCTGATGACGGGAGAAAAAGATAAAGTCGATGAAGACGGGA 654
Qy 833 TAAAGAACTAGAAAAGATCTGATGAGAGAAATTAACGCTGTGAAAAATCAACGTG 892
Db 655 AAAGACCGATCGAAAAAGATCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCAAGG 714
Qy 893 TTGATATTTTAGACGCTGACTTATGGGCAAGATCACGTTGTTTAATTTCTTCAGAGAA 952
Db 715 TTATGAAGCGGACAGATTCGCTGCTCGCTCCGACGCGCAGTTCGACGAGCAGATGAGC 774
Qy 953 TGGCAAAAGAACTTTGGA 971
Db 775 TGATGATGAATCGATCA 793

RESULT 23

ACA21545
ID ACA21545 standard; DNA; 972 BP.

XX AC ACA21545;

XX AC ACA21545;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #3202.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Bacillus anthracis.

OS WO200277183-A2.

XX WO200277183-A2.

XX 03-OCT-2002.

XX

21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
P-ESDB; ABU17675.
New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
Claim 14; SEQ ID NO 9415; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 972 BP; 405 A; 147 C; 183 G; 237 T; 0 U; 0 Other;

Query Match 14.2%; Score 143.2; DB 8; Length 972;

Best Local Similarity 55.0%; Pred. No. 4.7e-17;

Matches 308; Conservative 0; Mismatches 243; Indels 9; Gaps 1;

Qy 25 AAAATTTTAAAGTGAATTCGGCTTATTTGTTTAAATTCGCACTGCGAGCATGTGGAAT 84
Db 7 AAAATTCAGTATTTTCATAGTAGTTTCTTATTCGCTGTGGATCGGACAGCAAAA 66
Qy 85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAAGATGGAGTTGAAATCAAGCAAG 144
Db 67 GAGGAGAAAAGAAACAAAAGCGGACATAAAAAATCAAGCTATAACAAATTAACACGCT 126
Qy 145 GAAGGTACTACGAAAGTACCTAAACACCTTAACGTTGTTGTTCTTGAGTATTCATTT 204
Db 127 GAAGGGGAAACGAAAGTTAGATATAACCCAGCGAAAAAGTAGTTGTACTTTGAATGGTAT 186
Qy 205 GTTCATGCGTTAGTTGCTTTTAGATGTTTAAACCTGTTCGGATAGCGGATGATAACAAAA 264

Db 187 TCAGAGACTTATTAGCCTTGGTGTTCAGCCAGTAGGATGGCAGACATTAAAGAAATTAT 246
QY 265 AATCGTATTATTAACCATTAAGAGATTAATAATGGAAATACACTTCTGTAGGAACACGT 324
Db 247 AATAAATGGGTAAATACAAAAACAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 306
QY 325 AAGCACTTAAGTAAAGAAATCAGTAAACCTTAAACCCAGATTTAAATTATTGCTGATAAT 384
Db 307 CAACAACTTAAGAGAAATTAAGCCGTTTAAACCCAGATTTAAATTATTATCACACTTCA 366
QY 385 AATAGACACAAAGGTATTATTAAGACTTAAATTAATAATTTGCTTACGAT-----T 435
Db 367 TTCCGTTGGTAAAGCAATTAATAATGAATTAAGAACAAATTCACCAACAGTTATGTTTGTAT 426
QY 436 GAACGTAAAGTTTCAGTGGAGATTAATGAATTAATTTGATGCTTTTAAACAAATTTC 495
Db 427 CCATCAACAGCAATAACGATCACATTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 486
QY 496 AAGCTTTAGGTAAAGAAAGAGAGTAAACAAACGCTTAGAAGAACACGATTAAGAAATTT 555
Db 487 AAGCAGTTGGAAAGAGAGAGAGTAAACAAACGCTTAGAAGAACACGATTAAGAAATTT 546
QY 556 GAAGAATATAAAAAAGAAAT 575
Db 547 GCTGATGCAAAAGCAAAAT 566
RESULT 24
ID AAS49530/c
XX AAS49530 standard; DNA; 321 BP.
AC AAS49530;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #754.
XX
XX Antisense; ss; prokaryotic cellular proliferation; antibiotic;
KW antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX
PS Claim 1; SEQ ID NO 2107; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia

CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence is an antisense oligonucleotide of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 321 BP; 83 A; 73 C; 36 G; 129 T; 0 U; 0 Other;
Query March 12.7%; Score 128; DB 4; Length 321;
Best Local Similarity 67.8%; Pred. No. 2.7e-14;
Matches 179; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 132 AATCAAGCACGAAGGCTACTACGAAAGTACCTAAACACACCCCTAAACGTTGTTGTTCT 191
Db 271 AATTAAGATGATTAAGGAACTGAAATAATTAAGAAATCTTAACGTTGTTGTTAT 212
QY 192 TGAGTATTCAATTTGTTGATGCTTAGTTGCTTTTAGATGTTAAACCTTTGGGATAGCGGA 251
Db 211 AGAATATAGTTTTCGCTGATTATTTAGCAGCATTAGATATGAACCTGTTGGTTCGAGA 152
QY 252 TGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAATTTGGAATACACTTC 311
Db 151 TGATGGCAGCACTAAAAATATAACAAAGCTCAGTAAGAGATAGATTGGGGCATATGAATC 92
QY 312 TGTAGGACACGTAAAGCAACCTTAACCTTAGAAGAAATCAGTAAACCTTAACACGATTTAAT 371
Db 91 GGTGGATCTAGACCGCAACCGAATATCGAAGTATAGTAATTAATTAACCGGATTGAT 32
QY 372 TATTGCTGATTAATATAGACACAA 395
Db 31 CATTGCAGATGTTAGCAGACATAA 8
RESULT 25
ACA16735/c
ID ACA16735 standard; DNA; 321 BP.
XX
AC ACA16735;
XX
DT 27-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene antisense oligonucleotide #4605.
DE Antisense; ss; prokaryotic essential gene; cell proliferation;
KW drug design.
XX
XX Archaea.
XX
XX WO200277183-A2.
PN
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX
XX Claim 1; SEQ ID NO 4605; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the 6213
XX antisense sequences of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
XX standardise OS field)

SQ Sequence 321 BP; 83 A; 73 C; 36 G; 129 T; 0 U; 0 Other;

Query Match 12.7%; Score 128; DB 8; Length 321;
Best Local Similarity 67.8%; Pred. No. 2.7e-14;
Matches 179; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 132 AATCAAGCAGGAGGAGGTACTACGAAGTACTTAACACCTTAACCTGTTGTCT 191
DB 271 AATAAAGATGAATTAGGAACCTGAATAAATTAAGAAATCTTAACCTGTTGTATT 212
QY 192 TGAGTATTTCATTGTTGATGCGTTAGTCTTGTAGTGTAACTGTTGGATAGCGGA 251
DB 211 AGAATAATAGTTTGTGATATTATTAGCAGCATTAGATATGAACCTGTTGTTGCAGA 152
QY 252 TGATAACAAAAAATCGTATTATTATTAACCATTAAGAGATAAATTTGAAATATACACTTC 311
DB 151 TGATGGCAGCACTAAATAATATAACAAAGTCAGTAAGAGATAAGATTGGGCATATGAATC 92
QY 312 TGTAGGAACAGTAAGCAACCTAATCTTAGAAGAAATCAGTAACCTTAACCGATTATAT 371
DB 91 GGTGTGATCTAGACCGCAACCGAATATGGAAGTATGAATTAATTAACCGGTTGAT 32
QY 372 TATTGCTGATATATAGACACAA 395
DB 31 CATTGCAGATGTTAGCAGACATAA 8

RESULT 26
ID AAV74915
XX AAV74915 standard; DNA; 2115 BP.

AC AAV74915;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #604.
XX
KW Computer readable medium; vaccine; *S. aureus* infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT misc_feature 541..600
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
XX
PN EP786519-A2.
XX
PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-00100117.
XX
PR 05-JAN-1996; 96US-0009861P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
PT stored on computer readable medium and used in the production of anti-
PT *S. aureus* vaccines.
XX
PS Claim 1; Page 1535-1536; 3271pp; English.
XX
CC This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the *S. aureus* DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against *S. aureus* infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the *S. aureus* DNA sequences contained on the computer
CC readable medium
XX
SQ Sequence 2115 BP; 787 A; 279 C; 363 G; 622 T; 0 U; 64 Other;
Query Match 12.4%; Score 125.4; DB 2; Length 2115;
Best Local Similarity 53.2%; Pred. No. 8.9e-14;
Matches 289; Conservative 1; Mismatches 247; Indels 6; Gaps 1;
QY 458 ATTATAATGAAAATATTGATGCTTTTAAACAAATTTCAAAAGCTTTAGTAAAGAAAG 517
DB 1 ATTATAATGCAATATATTGAGCAATTTAAACAGTCGCTTAAGCAGTAGCAAGAAAG 60
QY 518 AAGGTAAAAACCGCTTAGAAGAACACGATAGAAAAATTTGAAGAAATATAAAAAGAAATAA 577
DB 61 AAGCGGAGAGCGCTCTGGAAAAAGCATGATATAAATATTAGCGGAGATTAGAAAAAATTG 120

QY 578 CTATGATTAATAATCAAAAGGTAATTCCTCGCAGTAGCTGCTAAATCAGGTTTGCTGTC 637
DB 121 AACAGAGTACGTTAAATCTGCATTTGCAATTCGGTATCTCAAGACGAGTATGTTTATTA 180
QY 638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAAAGAACGATTA 697
DB 181 ATAATGAGATACATTTATGCGACAACTCTTAATTAATGATGGTATTCACCTGAAGTCA 240
QY 698 GTGATGATGTTACTAAGGTTTAAAGTATCTTAAGTATCTTAAGGACCTTACTTACAATGAACA 757
DB 241 MAARAARAAAACCTACGCATGTTGGTGAACGCAAGGGTGGCTCTTATATATATTTAAATA 300
QY 758 CTGAATCTTATCTCAAGTGAATCTCTGAGCGTATGTTCAATATGACAAAACAAGCAAGTT 817
DB 301 ATGAAGACTTGGCAATATCAATCCAAAAGTATGATTTTATGCTGACGGAAGAACCG 360
QY 818 CTAACGAACCTTCACTAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAACGCTG 877
DB 361 ACAAAA-----TAGAACGAAATTCATTGATCCTGACGTTTGGAAATCATTAAAAAGCTG 414
QY 878 TGAATAATCAAGTGTGATATTTAGACCGTGACTTATGCGCAAGATCAGCTGTTTAA 937
DB 415 TGAAGATACAAAGTTTATGACGTTGACCGAAATTAAGTGGTTGAAATCAAGGGGATTA 474
QY 938 TTCTTCAGAGAAATGCGAAAGAACTGTTGGAATTTATCTAAGAAAGATAGTAAAAAG 997
DB 475 TCGCAAGTGAAGTATGGCAGAGATTTAGAAAAATTCGAGAAAAGCAAAATANAAT 534
QY 998 ATA 1000
DB 535 ACA 537

RESULT 27
ABT15032
ID ABT15032 standard; DNA; 990 BP.
XX
AC ABT15032;
XX
DT 06-MAR-2003 (first entry)
XX
DE Pathogen specific antigen related staphylococcal DNA SEQ ID No 318.
XX
KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis; gene; ds.
XX
OS Staphylococcus sp.
XX
PN W0200259148-A2.
XX
PD 01-AUG-2002.
XX
PF 21-JAN-2002; 2002WO-EP000546.
XX
PR 26-JAN-2001; 2001AT-00000130.
XX
(CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI Meinke A, Nagy E, Von Ahnen U, Klade C, Henics T, Zauner W;
PI Manh DB, Vytvytska O, Etsz H, Dryla A, Weichhart T, Hafner M;
PI Tempelmaier B;
XX
XX WPI; 2003-075410/07.
XX
XX Identifying, isolating and producing hyperimmune serum-reactive antigens
XX from a pathogen, for preparing vaccine or medicament for treating or
XX preventing e.g. staphylococcal infections, comprises providing antibody
XX preparation.
XX
XX Example 7; Page 210; 252pp; English.

XX The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens are
CC used in a vaccine, comprises providing antibody preparation from a plasma
CC pool of a type of animal, or individual sera with antibodies against the
CC specific pathogen, tumour, allergen, tissue or host prone to auto-
CC immunity. The hyperimmune serum-reactive antigens comprising any of the
CC 62 sequences of 53-2261 amino acids fully defined in the specification,
CC or their hyperimmune fragments are useful for the manufacture of a
CC pharmaceutical preparation, particularly a vaccine against staphylococcal
CC infections or colonisation against *S. aureus* or *S. epidermidis*. The
CC preparation of antibodies is useful for the manufacture of a medicament
CC for treating or preventing staphylococcal infections or colonisation
CC against *S. aureus* or *S. epidermidis*. The antibody preparations may also
CC be used for diagnostic and imaging purposes. Other conditions that can be
CC treated include cancer, autoimmune diseases or infections caused by viral
CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC polynucleotide sequence represents staphylococcal DNA relating to the
CC method for identifying and producing pathogen specific antigens of the
CC invention
XX
SQ Sequence 990 BP; 398 A; 141 C; 179 G; 272 T; 0 U; 0 Other;
Query Match 12.3%; Score 124; DB 8; Length 990;
Best Local Similarity 50.7%; Pred. No. 1.5e-13;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
QY 27 AATTTAAGTGAATGCTTATTTGTTTGTGTTTAACTGCACCTGCAGCATGCGAAATAA 86
DB 12 AATTAATAATGCTTGTGTTTACGCTTGTCTTCTTCTTTTGTAGCAGGATGTAGTGGAA 71
QY 87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
DB 72 TTCAATAAACAATCATCTGATCAAAAGATGAAGAAACACTTCAATTAACATGCAAT 131
QY 147 AGGTACTACGAAAGTACCTAAACACCCCTAAACGCTGTTGTTCTTGTAGTATTCATTTGT 206
DB 132 GGGTACAACTGAAATTAAGGGGAAACCAAGCGTGTGTTACGCTATATCAAGGTGCCAC 191
QY 207 TGATCGGTAGTGTCTTTAGATGTTTAAACCTGTTGGATAGCGGATGATAACAAAAAAAA 266
DB 192 TGACCTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAAATCATATGACACA 251
QY 267 TCGTATTATTAAACCATTAAGAGATATAAATTCGAAAATACACTTCTGTAGGAACACGTA 326
DB 252 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATACTAGATTTGTAGTCAAGAAC 311
QY 327 GCAACCTACTTAGAAGAAATCAGTAAACTTAAACAGATTTAAATTTATTTGCTGATATA 386
DB 312 TGCACCTACTTAGAGGAATCTCTAAATTAAGACCGGACTTAATTTGTCGCTCAAAAGT 371
QY 387 TAGACACAAGGTATTTATTAAGACTTAATAAATTTGCTCTCTACGATTTGAACCTGAAG 446
DB 372 TAGAAATGAAAAAAGTTTACGATCAATTTATCTAAATTCGACCC-----AACAG 418
QY 447 TTTTCGATGAGATTTAATGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
DB 419 TTTCTACTGATACAGTTTCAATTTCAAGATACAACTAAGTTAATGGGAAAGCTTTAG 478
QY 506 GTAAAGAAGAAGAGTAAAAACCGCTTAGAAGAACACGATAGAAAAATTTGAAGATATA 565
DB 479 GGAAGAAAAAGAGCTGAAGATTACTTTAAAGATGACGATGATAAAGTAGCTGCTATCC 538
QY 566 AAAAAA-----AATACTATGGATAAAAATCAAAAGGTATTTGCTGACGATGCTGCTAAT 622
DB 539 AAAAAAGATCAAAAGCAAGATTAAGATGCGCCATTGAAAGCTTCAGTTGTTAACT 598
QY 623 CAGGTTTGTCTGCTCATCAAGCAACTCTTTATGTTGCTCAATTTCTCAAGTCAACTAGGTT 682
DB 599 TCCGT---GCTGATCATACAAGAAATTTATGCTGGTGGATATGCTGGTGAATCTTAATG 655
QY 683 TTAAGAAGCAATTAAGTGATGATGTTTACTAAAGGTTTAAAGTAAATCTTTAAAGGACCTT 742

Db 656 ATTGAGGATCAAGCTAATTAAGACTTACAAAACAAGTTGATTAATCGTAAAGATATTA 715
Qy 743 ACTTACAAATGAACACTGAATCTTATCTCAAGTGAATCCTGAGCGTATGTTCAATGA 802
Db 716 TCCAACTTACATCTAAAGAAAGCATTCCTAATTAAGAACGCTGATCATATTTTGTAGTAA 775
Qy 803 CAACAAGAGCAAGTTCTTAACGAACCTTCACCTA 834
Db 776 AATCAGATCCAAATGCGAAGATGCTGCATTA 807

RESULT 28

AAZ22850
ID AAZ22850 standard; DNA; 999 BP.

XX AAZ22850;

AC 20-DEC-1999 (first entry)

DT Staphylococcus aureus cbrA DNA.

XX Infection; detection; diagnosis; screening; antibiotic; resistance;
KW methicillin; MRSA; ds.

XX Staphylococcus aureus.

OS Key Location/Qualifiers

XX 7..999

FT /*tag= a

FT /product= "S. aureus cbrA protein"

XX W09947662-A1.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US006199.

XX 20-MAR-1998; 98US-0078682P.

PR 01-APR-1998; 98US-0080296P.

PR 07-MAY-1998; 98US-0084674P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (LUDW-) LUDWIG INST CANCER RES.

XX Simpson AJG, Choi GH;

XX WPI, 1999-580304/49.

XX P-PSDB; AAY42331.

XX Staphylococcus aureus genes and polypeptides, vectors and methods of
PT recombinant production.

XX Claim 1; Page 24; 111pp; English.

XX This sequence represents Staphylococcus aureus cbrA DNA. The cbrA protein
CC is predicted to have a molecular weight of about 36.8 kD and, along with
CC cbrB (AAY42332) and cbrC (AAY42333), is thought to be involved in iron
CC regulation, based on amino acid sequence homology with known iron
CC regulator proteins. S. aureus is a ubiquitous pathogen which causes
CC infections in burns, cellulitis, eyelid infections, food poisoning, joint
CC infections, neonatal conjunctivitis, osteomyelitis, skin infections,
CC surgical wound infection, scalded skin syndrome and toxic shock syndrome.
CC S. aureus is increasingly becoming resistant to known antibiotics, with
CC methicillin-resistant strains generally being multiply drug resistant.
CC Methicillin-resistant S. aureus (MRSA) poses serious infection control
CC problems, with many strains being multiresistant against virtually all
CC antibiotics with the exception of the vancomycin-type glycopeptide
CC antibiotics. The protein may be useful to screen potential antipeptidase
CC which could be used as antibiotics and it may be used as a vaccine to
CC prevent or attenuate an infection caused by a member of the
CC Staphylococcus genus. The protein, or antibodies against it can be used
CC in immunoassays to detect Staphylococcus in a biological sample. Probes

CC and primers derived from the nucleic acid sequences may also be used to
CC detect Staphylococcus nucleic acids in a biological sample

XX SQ Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

Query Match 12.3%; Score 124; DB 2; Length 999;

Best Local Similarity 50.7%; Pred. No. 1.5e-13;

Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

Qy 27 AATTTAAAGTAAATGGCTTATTTGTTTAAATGCAACTGAGCATGTGGAATAA 86

Db 18 AATTTAAATGCTTGTGTTACGCTTGCTTCTACTTGTCTTTAGCAGGATGTAGTGGAA 77

Qy 87 TAGTTCAAGTAACTCAAGTAAAGATCATCAAAAGATGGAGTTGAAATCAAGCAGAGA 146

Db 78 TTCAATAAACAATCATCTGTATAACAAGATAAGGAACAACCTTCAATTAACATGCAAT 137

Qy 147 AGGTACTACGAAAGTACCTTAAACACCTTAAACGTTGTTGTTCTTGTAGTATTTCATTTGT 206

Db 138 GGGTCAACTGAATTAAGGGAAACCAAGCGTGTGTTACGCTATATCAAGGTGCCAC 197

Qy 207 TGATGCGTTAGTCTTTTAGATTTAAACCTGTTGGGATAGCGGATGATTAACAAAAA 266

Db 198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGTGTCTGTAGAAATCATGGACACAAA 257

Qy 267 TCGTATTATTAAACCATTAAGAGATAAAATTTGGAATATACACTTCTGTAGGAACACGTAA 326

Db 258 ACCGAAATTCGAATACATATAAAATGATTTTAAAGATACCTAAGATTTAGTCAAGAAC 317

Qy 327 GCAACCTTAACCTTAGAAGAAATCAGTAAACTTAAACCCAGATTTAAATTTGCTGATAATA 386

Db 318 TGCACCTTAACCTTAGAGGAAATCTCTAAATTAACCCGACCTTAATTTGCGGTCAAAAGT 377

Qy 387 TAGACACAAAGGTATTATTAAGACTTAAATAAATTTGCTCTACGATTGAACGTGAAAAG 446

Db 378 TAGAAATGAAAAAGTTTACGATCAATTTATATATATATATATATATATATATATATATAT 424

Qy 447 TTTCGATGAGATTTATTAATGAATAT-ATTGATGCTTTTAAACCAATTTCAAAAGCTTTAG 505

Db 425 TTTCCTACTGATACAGTTTTCAAATTTCAAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 484

Qy 506 GTAAAGAGAAAGAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAAATTTGAAGAATAATA 565

Db 485 GGAAAGAAAGAGAGCTCAAGATTTTACTTAAAGATACGATGATAAAGTAGCTGCATTC 544

Qy 566 AAAAGA---AATACTATGATATAAATCAAAAGGTATTGCTGCGAGTAGCTGCTAAAT 622

Db 545 AAAAGATGCAAAAGCAAAAGTATAAGATGATGCCATTGAAAAGCTTCAGTTGTAACT 604

Qy 623 CAGGTTTGTGCTCATCCCAAGCACTTTATGTTGGTCAATTCCTAAGTCAACTAGCTT 682

Db 605 TCCGT---GCTGATCATACAAAGATTTATGCTGGTGGATATGCTGGTGAATCTTTAAATG 661

Qy 683 TTAAAGAGCAATTAAGTATGATGATGTTACTAAAGGTTTAAAGTAAAGTATCTTAAAGGACCTT 742

Db 662 ATTTAGGATTCAAAGTAAATAAAGACTTACAAAACAAGTTGATAATGGTAAAGATATA 721

Qy 743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATTCCTGAGCGTATGTTTCAATGA 802

Db 722 TCCAACCTTACATCTAAAGAAAGCATTCCTAATTAAGAACGCTGATCATATTTTTGTAGTAA 781

Qy 803 CAACAAGCAAGTTCTTAACGAACCTTCACCTA 834

Db 782 AATCAGATCCAAATGCGAAGATGCTGCATTA 813

RESULT 29

AAZ19889

ID AAZ19889 standard; DNA; 999 BP.

XX AAZ19889;

XX 06-DEC-1999 (first entry)

DT

XX Staphylococcus aureus iron regulation gene cbrA.
 DE CbrA gene; infection; therapy; diagnosis; vaccine; antibiotic;
 KW iron regulation; ss.
 KW Staphylococcus aureus.
 OS
 XX
 PH Key Location/Qualifiers
 FT CDS 7..999
 FT /*tag= a
 XX
 XX WO947639-A2.
 XX
 XX
 XX 23-SEP-1999.
 XX
 XX 19-MAR-1999; 99WO-US005976.
 XX
 XX 20-MAR-1998; 98US-0078682P.
 PR 01-APR-1998; 98US-0080296P.
 PR 07-MAY-1998; 98US-0084574P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Simpson AJG, Choi GH;
 XX
 XX WPI; 1999-562101/47.
 DR P-PSDB; AAY31824.
 XX
 XX New isolated Staphylococcus nucleic acid molecules, used to develop
 PT products for the diagnosis, prevention and treatment of Staphylococcal
 PT infections.
 XX
 XX Claim 1; Page 24; 102pp; English.
 XX
 XX This is the nucleotide sequence of Staphylococcus aureus strain ISP3
 CC (ATCC 202108) genomic DNA including the novel cbrA gene that codes for a
 CC 330-amino acid protein (see AAY31824) of predicted mol.wt. 36.8 kDa. The
 CC sequence was obtained from overlapping clones BTAC44 and BTAGJ54, which
 CC span a single operon containing the cbrA, cbrB and cbrC genes (see
 CC AAZ19889-91). CbrA shows sequence homology to known genes involved in
 CC iron regulation. The invention provides 11 novel genes (see AAZ19882-92)
 CC of S. aureus and the polypeptides they encode (see AAY31817-27). Also
 CC provided are vectors, host cells, antibodies and hybridomas. The
 CC invention further relates to screening methods for identifying agonists
 CC and antagonists of S. aureus polypeptide activity, and to diagnostic
 CC methods for detecting Staphylococcus nucleic acids, polypeptides and
 CC antibodies in a biological sample. Antagonists of cbrA may be useful as
 CC antibiotics to treat infections of S. aureus and other Staphylococcus spp.
 CC Also provided are novel vaccines for the prevention or attenuation of
 CC infection by Staphylococcus. The isolated nucleic acid molecule is also
 CC useful for generating probes and primers, and in the recombinant
 CC production of cbrA protein
 XX
 XX Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;
 Query Match 12.3%; Score 124; DB 2; Length 999;
 Best Local Similarity 50.7%; Pred. No. 1.5e-13;
 Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
 27 AATTTTAAAGTAAATGGCTATTGTTGTTTAAATGCAACTGCAGCATGTGGAATAA 86
 18 AATTTAAATGCTGTTGTTGTTGCTTCTCTACTGTTGTTTACGAGATGTAGTGGAA 77
 87 TAGTTCAGTAACCTCAAGTAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
 78 TTCAATTAACCAATCACTGTATACAAAGATAAGGAACAACACTTCAATTAAACATGCAAT 137
 147 AGGTACTACGAAGTACCTAAACACCTTAAACCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 206
 138 GGGTACAACTCAAAATTAAGGGGAACCAAGCGTGTGTTACGCTATATCAAGGTGCCAC 197

QY 207 TGATCGGTTAGTGGCTTTAGATGTTAAACCTGTTGGATAGCGGATGATAACAAAAA 266
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 257
 198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAAATCATGGACAAAAA
 QY 267 TCGTATTATTAAACCATTAAGAGATAAAATTTGGAATAATACACTTCTGTAGGAACACGTAA 326
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 317
 258 ACCGAAATTCGAATACATATAAAATGATTTAAAGATACTAAGATTGTAGGTCAAGAAC 317
 QY 327 GCAACCTAACTTAGAAGAAATCAGTAAATCTTAAACCCAGATTAAATTTGCTGTAATAA 386
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 377
 318 TGCACCTAACTTAGAGGAAATCTCTAAATTTAAACCCGACTTAATTTGCGGCTCAAAAGT
 QY 387 TAGACAAAAGGTATTATTAAGACTTAAATTAATAATGCTCCTAGATTGAATGAAAAG 446
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 424
 378 TAGAAATGAAAAGTTTACGATCAATTTATCTAAATTCGCACC-----AACAG
 QY 447 TTTCGATCGAGATTATATGAAAT-ATTGATGCTTTTAAACCAATTTCAAAAGCTTTAG 505
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 484
 425 TTTCCTCTGATACAGTTTCAATTTCAAGATACAACTAAGTTATGGGAAGCTTTAG
 QY 506 GTAAAGAAAGAGAGGTAAAAACGCTTTAGAGAAACACGATAAGAAAAATTTGAAGAATA 565
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 544
 485 GGAAGAAAGAGAGCTGAAGATTCTTAAAGAGTACGATGATAAGTAGCTGCAATCC
 QY 566 AAAAAA---AATAACTATGGATAAAATCAAAAGGTATTGCTCGCTAGCTAGCTTAAT 622
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 604
 545 AAAAAAGATGCAAAAGCAAGTATAAAGATGCAATGCGCATTTGAAAGCTTCAGTTGTTAACT
 QY 623 CAGGTTTGGCTGCTCATCCAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 661
 605 TCCGT---GCTGATCATCAAGAAATTTATGCTGGTGGATATGCTGGTGAATCTTAAATG
 QY 683 TTAAAGAAAGCATTAAGTGTATGTTACTTAAAGGTTTAAAGTAAATCTTAAAGGACCTT 742
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 721
 662 ATTTAGGATTCAAACGTAATAAGACTTTACAAAACAAAGTTGATAATGCTAAAGATATTA
 QY 743 ACTTCAAAATGAACACTGAACTTTATCTCAAGTGAATCTCGAGCGTATGTTCAATAA 802
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 781
 722 TCCAACTTACATCTAAAGAAAGCAATTCCTCAATTAATGAAGCTGATCATATTTTTGTAGTAA
 QY 803 GAAACAAAGCAAGTTCTAACGACCTTCACTA 834
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 813
 782 AATCAGATCCAAATCGAAAGATGCTGCAATTA
 RESULT 30
 ADF43555
 ID ADF43555 standard; DNA; 999 BP.
 XX
 AC ADF43555;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Staphylococcus aureus cbrA DNA.
 XX
 KW antibacterial; immunostimulant; vaccine; vaccine;
 KW Staphylococcus aureus infection; infection prevention;
 KW infection attenuation; gene; ds; cbrA.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT CDS 7..999
 FT /*tag= a
 FT /product= "cbrA"
 XX
 XX US2003153733-A1.
 PN 14-AUG-2003.
 XX
 XX 24-OCT-2002; 2002US-00278946.
 XX
 XX 20-MAR-1998; 98US-0078862P.


```

PR 01-APR-1998; 98US-0080296P.
PR 07-MAY-1998; 98US-0084674P.
PR 18-MAR-1999; 99WO-US006199.
PR 15-JAN-2002; 2002US-00830217.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Simpson AJG, Choi GH;
XX
XX WPI; 2003-897737/82.
DR P-PSDB; ADF43556.
DR
XX New Staphylococcus aureus gene, useful for preparing a vaccine against
PT infection caused by Staphylococcus aureus.
XX
XX Claim 1; SEQ ID NO 15; 61pp; English.
XX
XX The invention describes a new isolated nucleic acid molecule comprising:
CC a sequence that is 95% identical to a fully defined sequence having 586-
CC 2226 bp, or its complement; or a sequence encoding a polypeptide having a
CC fully defined sequence comprising 136-691 amino acids, or its complement.
CC The gene is Staphylococcus aureus gene. The nucleic acid is useful for
CC preparing a vaccine against infection caused by Staphylococcus aureus.
CC The methods are useful for preventing or attenuating an infection caused
CC by a Staphylococcus, detecting Staphylococcus nucleic acids in a
CC biological sample obtained from an animal, and detecting Staphylococcus
CC antibodies in a biological sample obtained from an animal. This sequence
CC encodes Staphylococcus aureus cbrA protein.
XX
SQ Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

Query Match 12.3%; Score 124; DB 10; Length 999;
Best Local Similarity 50.7%; Pred. No. 1.5e-13;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

QY 27 AATTGTAAGTGAATGGCTTATTGTTGTTTAAATGCACTGCAGCATGTGGAATAA 86
DB |||||
18 AATTAATGCTGTGTGTAGCTTCTTCTACTGTTTGTAGCAGGATGTAGTGGAA 77
QY 87 TAGTTCAGTAACCTCAAGTAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
DB |||||
78 TTCAAAATAAACAACTCATCTGATAACAAGATAAGGAACAACTTCAATTAAACATGCAAT 137
QY 147 AGGTACTAGGAAGTACCTAAACACCTTAAGCTGTTGTTGTTCTTGAGTATTCAATTGT 206
DB |||||
138 GGGTACAACCTGAATTTAAAGGGAACCAAGCGTGTGTTAGCTTATATCAAGGTGCCAC 197
QY 207 TGATCGGTGTAGTGTGTTTGTAGTGTAAACCTGTTGGGATAGCGGATGATAACAAAAAAA 266
DB |||||
198 TGACGTGCTGTATCTTTTGTAGTGTAAACCTGTTAGCTGTAGTGTAGTGTCAAGAAC 257
QY 267 TCGTATTATTAACCAATTAAGAGATAAAATTTGGAATAATACATCTCTAGGAACACGTA 326
DB |||||
258 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATATAAGATTGTAGGTCAAGAAC 317
QY 327 GCAACTAATCTAGGAAGAAACAGTAACTTAACCAAGATTATTAATTTGCTGATAATA 386
DB |||||
318 TGCACCTAATCTAGGAAGAAATCTTAAATTTAAACCCGGAATTAATTTGCGGTCAAAAGT 377
QY 387 TAGACACAAAGGTATTTTAAAGACTTAAATAAAATGCTCTCTACGATTGAACCTAAAAAG 446
DB |||||
378 TAGAATGAAAAGTTTACGATCAATTAATCTTAAATTCGACC-----AACAG 424
QY 447 TTTCGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAACAATTTCAAAAGCTTTAG 505
DB |||||
425 TTCTACTGATACAGTCTTCAAAATCAAAGTAACTAAAGTTAATGGGAAAGCTTTAG 484
QY 506 GTAAGAGACAGAGGTAAAAACCTTAGAGACACAGTAAGAAATGAGATATA 565
DB |||||
485 GGAAAGAAAAGAGCTGAAGATTTCATTAAGAGTACGATGAAGTAGTGCATTC 544
QY 566 AAAAAGA---AATAACTATGGATAAAATCAAAAGGTATTGCTCGAGTAGCTGCTAAAT 622
|||

```

```

Db 545 AAAAGATGCAAAAGCAAAAGTATAAAGATGCATGGCCATTGAAAAGCTTCAGTTGTTAACT 604
QY 623 CAGGTTTGTCTGCTCATCCAAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGTT 682
Db |||||
605 TCCGT--GCTGATCATACAGAAATTTATGCTGGTGGATATGCTGGTGAATCTTAAATG 661
QY 683 TTAAAGAGCAATTAAAGTGATGATGTTACTAAAGGTTTAAAGTAACTATCTTAAAGGACCTT 742
Db |||||
662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAACAAAGTTGATAATGGTAAAGATATTA 721
QY 743 ACTTACAAATGAACACACTGAAACTTTTATCTCAAGTGAATCTCTAGCGTATGTTCAATAATGA 802
Db |||||
722 TCCAACCTTACATCTAAAGAAAGCAATTCATTTAATGAACGCTGATCATATTTTGTAGTAA 781
QY 803 CAAACAAAGCAAGTTCTTAACGAACCTTCACTA 834
Db |||||
782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813

RESULT 31
ACC48531/c
ID ACC48531 standard; DNA; 2940 BP.
AC ACC48531;
XX
DT 11-AUG-2003 (first entry)
XX
DE Staphylococcal surface-exposed immunogenic polypeptide DNA.
XX
KW Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW antibacterial; vaccine; gene; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS complement (998..1990)
FT /*tag= a
FT /product= "SEIP"
XX
PN WO2003020875-A2.
XX
PD 13-MAR-2003.
XX
PF 17-JUN-2002; 2002WO-US019224.
XX
PR 17-JUN-2001; 2001US-0298975P.
XX
PA (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI Scott DL;
XX
DR WPI; 2003-300870/29.
XX P-PSDB; ABR41864.
XX
PT Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
PT containing receptors for siderophores or iron-binding ligands, useful for
PT producing antibodies effective against Staphylococci infection.
XX
PS Claim 4; Page 51-52; 62pp; English.
XX
CC The present sequence is that of DNA encoding novel Staphylococcus aureus
CC surface-exposed immunogenic polypeptide (SEIP) D2 SA02. To isolate SEIP
CC genes, an expression library of S. aureus genomic DNA was screened with
CC anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
CC amino acid sequences were deduced and analysed for conserved and/or
CC functional domains. The SEIP encoded by the present DNA sequence has
CC sequence homology with the siderophore family of periplasmic binding
CC proteins. The invention provides methods for the identification,
CC production and recovery of SEIPs. The SEIPs can be used individually, or
CC in combination, to produce anti-staphylococcal antibodies useful in
CC passive or active immunisation strategies to prevent or contain
CC staphylococcal infection. They can also be used to develop diagnostic
CC assays

```

XX SQ Sequence 2940 BP; 955 A; 514 C; 512 G; 959 T; 0 U; 0 Other;
Query Match 12.3%; Score 124; DB 8; Length 2940;
Best Local Similarity 50.7%; Pred. No. 1.6e-13;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
QY 27 AATTTTAAGTAAATGGCTTATTTGTTTAAATGCAACTGCAGCATGTGGAATAA 86
DB |||||
1979 AATTAATGCTTGTGTACGCTGCTTCTACTTGTGTAGCAGGATGTAGTGGAA 1920
QY 87 TAGTTCAAGTAATCAAGTAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGA 146
DB |||||
1919 TTAATAAACAATCATCTGTAACAAGATAAGGAACAATCAATTAATAACATGCAAT 1860
QY 147 AGGTACTAGAAAGTACCTAAACACCCCTAAACGCTGTTGTTCTGTAGTATTCATTGT 206
DB |||||
1859 GGGTACAACCTGAAATTAAGGGGAAACCAAGGCTGTTGTACGCTATATCAAGGTGCCAC 1800
QY 207 TGATCGTGTAGTTGCTTTAGATGTTAAACCTGTTGGATAGCGGATGATAACAAAAAAA 266
DB |||||
1799 TGACGTCGCTGTATCTTTAGTGTGTTAAACCTGTAGTGTGTAGTATCATGACACAAA 1740
QY 267 TCGTATTTAATTAACCAATTAAGAGATAAAATTTGAAATATACCTTCTGTAGGAACAGTAA 326
DB |||||
1739 ACCGAATTCGAATACATATAAAATGATTTAAAGATATACTAAGATTGTAGGTCAAGAAC 1680
QY 327 GCACCTAATCTAGAAGAATCAAGTAACCTTAACCCAGATTTAATTTGCTGTAGTAATA 386
DB |||||
1679 TGACCTAATCTAGAGGAATCTTAATTTAAACCCGACTTAATTTGTCGCTCAAAAGT 1620
QY 387 TAGACAAAGGTATTTAATAGACTTAATAAATTTGCTTACGATTGAAGTGAAG 446
DB |||||
1619 TAGAAATGAAAGTTACGATCAATATCTAATATCGACCC-----AACAG 1573
QY 447 TTTTCGATGGAGATTATAAGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
DB |||||
1572 TTTCTACTGATACAGTTTTCATATTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 1513
QY 506 GTHAAGAAGAGGTAAAGCGCTTAGAAGACACAGTAAAGAAATTTGAAGATATA 565
DB |||||
1512 GGAAGAAGAAAGCTGAAGATTCTTAAAGATACGATGATAAAGTAGCTGCATCC 1453
QY 566 AAAAGA--AATACTATGATAAAATCAAAAGGTATTGCTGAGTAGCTGCTAAAT 622
DB |||||
1452 AAAAAGATGCAAAAGCAAGTATAAAGATGATGATGCGCATTTGAAAGCTTCAGTTTAACT 1393
QY 623 CAGGTTTGTGCTCATCCAGCAACTCTTATGTTGTCATTTCTTAAGTCAACTAGGTT 682
DB |||||
1392 TCCGT---GCTGATCATACAGAAATTTATGCTGTGATATGCTGTAATCTTAATG 1336
QY 683 TTAAGAAGCATTTAAGTATGATGTTATCAAAAGTTTAAAGTATGTTCTTAAGAGACCTT 742
DB |||||
1335 ATTTAGATTCAAAAGCTTAATAAGACTTACAAAACAAAGTTGATAATGTTAAAGATATTA 1276
QY 743 ACTTACAATGACACTGAACTTTATCTCAAGTATCTGAGCTGATGTTCTCATATGA 802
DB |||||
1275 TCCAACCTCATCTTAAGAAGAGCATTTCCATTAATGAACGCTGATCATATTTTGTAGTAA 1216
QY 803 CAACAAGCAAGTTCTTAACGAACCTTCACTA 834
DB |||||
1215 AATCAGATCCAAATGCAAGAGATGCTGCATTA 1184

RESULT 32

AAV74549

ID AAV74549 standard; DNA; 3775 BP.

XX AC AAV74549;

XX AC

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #238.

XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome; ds.
OS Staphylococcus aureus.
XX Key Location/Qualifiers
PH 1561..1620
FT /*tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
FT 3361..3420
FT /*tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
XX EP786519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-00100117.
XX 05-JAN-1996; 96US-0009861P.
XX (HUNA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX Claim 1; Page 1021-1023; 3271pp; English.
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX homologues of any of the S.aureus DNA sequences contained on the computer
XX readable medium
XX Sequence 3775 BP; 1222 A; 597 C; 633 G; 1199 T; 0 U; 124 Other;
SQ
Query Match 12.3%; Score 124; DB 2; Length 3775;
Best Local Similarity 50.7%; Pred. No. 1.7e-13;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
QY 27 AATTTTAAGTAAATGGCTTATTTGTTTAAATGCAACTGCAGCATGTGGAATAA 86
DB |||||
60 AATTAATGCTTGTGTACGCTTCTACTTGTGTAGCAGGATGTAGTGGAA 119
QY 87 TAGTTCAAGTAACTCAAGTAAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGA 146

Db 120 TTCAATTAACAACTCATCTGATACAAAGATAAGGAACAACTTCAATTAACATGCAT 179
Qy 147 AGGTACTACGAAGTACCTAAACACCCCTAAACGCTGTTGTTCTTGTAGTATTCATTTGT 206
Db 180 GGGTACAACTGAAATTAAGGGAACCAACGCTGTTGTTAGCTATATCAAGGTGCCAC 239
Qy 207 TGATCGTGTAGTGTGTTAGATGTTAAACCTGTTGGGATAGCGGATGATACAAAAA 266
Db 240 TGACGTCTGTATCTTTAGTGTGTTAAACCTGTTAGTGTCTAGAGTATCATGGACACAAA 299
Qy 267 TCGTATTATTAAACCATTAAGAGATAAAATTTGGAATATACACTTCTGTAGGAACAGTAA 326
Db 300 ACCGAAATTCGATACATAAAATGATTTAAAGATACAAAGTTGTAGTCAAGACC 359
Qy 327 GCAACCTAACTTAGAAGAAATCAGTAAATTAACACGATTTAAATTTATGCTGTAATAA 386
Db 360 TGCACCTAACTTAGAGGAATCTCTAAATTAACACCGGACTTAATTTGCGGTCAAAAGT 419
Qy 387 TAGACACAAAGGTATTATTAAGACTTAAATAAATTTGCTCTACGATTGAACGAAAG 446
Db 420 TAGAAATGAAAGGTTTACGATCAATTTATTAATAATCGACC-----AACAG 466
Qy 447 TTTCGATCGAGTTATTAATGAAAT-ATTGATGCTTTTAAACAACTTTCAAAAGCTTTAG 505
Db 467 TTTCTACTGATACAGTTTCAATTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 526
Qy 506 GTAAAGAAAGAAAGGTAAAAACGCTTAGAAGAACACGATGAAGAAATTAAGAAATATA 565
Db 527 GGAAGAAAGAAAGAGCTCAAGATTTTACITTAAGAAAGTACGATGAAGATAGCTGCAATCC 586
Qy 566 AAAAGA---AATACTATGATATAAATCAAAAGTATTGCTGCACTAGCTGCTAAAT 622
Db 587 AAAAGATGCAAAAGCAAGTATAAAGATGATGCGCCATTAAGAAAGCTTCAAGTTGTTAACT 646
Qy 623 CAGGTTTCTGCTCATCCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGCTT 682
Db 647 TCCGT---GCTGATCATACAGAAATTTATGCTGGTATATGCTGGTGAATCTTAATG 703
Qy 683 TTAAGAAAGCATTAAGTATGATGTTAATAAGGTTTAAAGTAAATCTTAAAGGACCTT 742
Db 704 ATTTAGGATTCAAGCTTAATAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763
Qy 743 ACTTACAAATGACACTGAATCTTATCTCAAGTGAATCCTGAGCGTATGTTCAATGA 802
Db 764 TCCAACTTACATCTTAAGAAAGCAATTCATTAATGAAGCTGATCATATTTTGTAGTAA 823
Qy 803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
Db 824 AATCAGATCCAAATCGAAGATGCTGCATTA 855

RESULT 33

ACA22139

ID ACA22139 standard; DNA; 796 BP.

XX ACA22139;

AC ACA22139;

XX 19-JUN-2003 (first entry)

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #3796.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX Bacillus anthracis.

OS WO200277183-A2.

XX 03-OCT-2002.

PD 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX

XX

XX

XX

XX

XX

XX

XX

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

P-PSDB; ABU18269.

New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to

isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 10009; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of

the 6213 antisense sequences given in the specification where expression

of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid

encoding a polypeptide whose expression is inhibited by the antisense

nucleic acid; (2) a host cell containing the vector; (3) an isolated

polypeptide or its fragment whose expression is inhibited by the

antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for

proliferation; (7) identifying a compound that influences the activity of

the gene product or that has an activity against a biological pathway

required for proliferation, or that inhibits cellular proliferation; (8)

identifying a gene required for cellular proliferation or the biological

pathway in which a proliferation-required gene or its gene product lies

or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a

product is overexpressed or underexpressed; (12) determining the extent

to which each of the strains is present in a culture or collection of

strains; or (13) identifying the target of a compound that inhibits the

proliferation of an organism. The antisense nucleic acids are useful for

identifying proteins or screening for homologous nucleic acids required

for cellular proliferation to isolate candidate molecules for rational

drug discovery programs, or for screening homologous nucleic acids

required for proliferation in cells other than *S. aureus*, *S. typhimurium*,*K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

prokaryotic essential genes. Note: The sequence data for this patent did

not form part of the printed specification, but was obtained in

electronic format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences

SQ

Sequence 796 BP; 324 A; 124 C; 146 G; 202 T; 0 U; 0 Other;

Query Match 11.2%; Score 113.2; DB 8; Length 796;

Best Local Similarity 57.6%; Pred. No. 1.4e-11;

Matches 227; Conservative 0; Mismatches 158; Indels 9; Gaps 1;

Qy 191 TTGATGATTCATTTGTTGATGCGGTAGTGTGTTAGATGTTAAACCTTTGGGATAGCGG 250

Db 2 TTGAATGGGTATATTTCAGAGACTTATTAGCACTTGTGTGTTTCAGCCAGTAGGGATGGCAG 61

Qy 251 ATGATACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTCGAAAAATACACTT 310

Db 62 ACATTAAGAATTATAAATGCGGTAAATACAAAAACAAACCGAGTAAAGATGTTGTAG 121

Qy 311 CTGTAGGAACACGTAAGCAACTTAACCTTAGAAGAAATCAGTAAATCTTAAACAGATTTAA 370

Db 122 ATGTCGGGACAGTCACACCAAACTTAGAAGAAATTAGCGGTTTAAACAGATTTAA 181

Qy 371 TTATTTGCTGATAATAATAGACACAAAGGTATTTTAAAGACTTTTAAATTTGCTCTTA 430

Db 182 TTATCATCAGCTTCATTCGTCGTAAGCAATTAATAAATGAATTAGAACAAATTCACCAA 241

QY 431 CGAT-----TGAACCTGAAAGTTTCGATGGAGTATATATGAAATATTGTGCTT 481
 Db 242 CAGTTATGTTTGTGTCATCAACAGCAATTAACGATCAGTTTGTGCTGAAATGACAGAAACAT 301
 QY 482 TTAACAACAATTTCAAAAGCTTTAGGTAAAGAGAGAGAGGTAAACCGCTTTAGAGAAC 541
 Db 302 TTAACAACAATTTCAAAAGCAGTTTGGAAAAAGAGAGAGGTAAACCGCTTTAGCTGATA 361
 QY 542 ACGATAAGAAAAATTGAAGAAATATAAAAAAGAAAT 575
 Db 362 TCGATAAGCAATTCGCTGATGCAAAAGCAAAAT 395

RESULT 34

ACF74610

ID ACF74610 standard; DNA; 861 BP.

XX ACF74610;

DT 20-NOV-2003 (first entry)

XX Staphylococcus aureus DNA #2290.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;

KW enzymatic assay; antibiotic target; gene; ds.

XX Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-18002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.

DR P-PSDB; ABM73050.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or

PT preventing Staphylococcal infection, specifically an infection caused by

PT S. aureus, e.g. sepsis.

XX Claim 6; SEQ ID NO 4579; 49pp; English.

XX The invention relates to novel genes and encoded proteins from

CC Staphylococcus aureus. A composition comprising the S. aureus protein, a

CC nucleic acid encoding the protein, or an antibody to the protein, is

CC useful as a pharmaceutical, particularly as a vaccine for treating or

CC preventing infection due to Staphylococcus bacteria, specifically an

CC infection caused by S. aureus. The composition is particularly useful for

CC treating or preventing sepsis in a patient. The composition can also be

CC used for diagnostics. The protein is also used in an assay for enzymatic

CC studies and as a target for antibiotics. This sequence represents one of

CC the novel S. aureus genes of the invention

XX Sequence 861 BP; 348 A; 123 C; 158 G; 232 T; 0 U; 0 Other;

SQ

Query Match 10.9%; Score 110.2; DB 8; Length 861;

Best Local Similarity 51.8%; Pred. No. 5.1e-11;

Matches 358; Conservative 0; Mismatches 313; Indels 20; Gaps 4;

QY 148 GGTACTACGAAGTACCTAAACACCTTAAACGGTGTGTTGTTCTTCAGTATTCATTGTT 207

Db 4 GGTACACTGAATTAAGGGAACCAACCGGTGTGTTACGCTATATCAAGTGCCACT 63

QY 208 GATCGCTGATGCTTTAGATGTAAACCTGTGGATAGCGGATGATACAAAAAAT 267

Db 64 GACGTCGCTGATCTTTAGGTGTTAAACCTGTAGTGTGCTGTAGATCATGGAACAAAA 123

QY 268 CGTATTATTAAACCATTAAGAGATAAAATTTGAAATACACCTTCTGTAGGAACACGTAAG 327
 Db 124 CCGAAATTCGATACATATAAAATGATTTTAAAAGATACCTAAGATTGTAGTCAAGAACCT 183
 QY 328 CAACCTACTTAGAAGAAATCAGTAAACCTTAAACAGATTTAAATTTATGCTGATATAAT 387
 Db 184 GCACCTAACTTAGAGGAAATCTCTAAATTTAAACCGGACTTAATTTGTGCGCTCAAAAGTT 243
 QY 388 AGACACAAAGGTATTATTAAAGACTTAAATAAAATTTGCTTACGATTGAACCTGAAAACT 447
 Db 244 AGAAATGAAAAGTTTACGATCAATTTATAAATTCGCACC-----AACAT 290
 QY 448 TTCTGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAACCAATTTCAAAAGCTTTAGG 506
 Db 291 TTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTTAAATGGGGAAGCTTTAGG 350
 QY 507 TAAAGAAGAGAGGTAAACCGCTTAGAAGAACAGATAGAAAAATTTGAAGATATAA 566
 Db 351 GAAAGAAAAAGAGCTGAAGATTACTTTAAAGATACGATGATAAAGTAGCTGCATTCCA 410
 QY 567 AAAAGA---AATAACTATGATAAAATCAAAAGGTATTGCTGAGTAGCTGCTAAAATC 623
 Db 411 AAAAGATGCAAAAGCAAAAGTATAAAGATGCGATGGCCATTTGAAAGCTTTCAGTTGTAACTT 470
 QY 624 AGGTTTGTCTCATCCAAAGCAACTTTATTTGTTGTCATTAATTCCTAAAGTCAACTAGTTT 683
 Db 471 CGGTGC---TGATCATACAGAATTTATGCTGCTGGATATGCTGGTGAATCTTAAATCA 527
 QY 684 TAAAGAAGCATTAAGTATGATGATGTTACTTAAAGTTTAAAGTATCTTAAAGGACCTTA 743
 Db 528 TTTAGGATTCAAACCGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTAT 587
 QY 744 CTTACAAATGAACACTGAACTTTATCTCAAGTGAATCCTGAGCGTATGTTTCATATGAC 803
 Db 588 CCACTTACATCTAAAGAAAGCAATTCCTAATAAGACGCTGATCATATTTTTGTAGTAAA 647
 QY 804 AAACAAAGCAAGTTCTTAACGAACCTTCACTA 834
 Db 648 ATCAGATCCAAATGCGAAAGATGCTGCATTA 678

RESULT 35

AAH54418

ID AAH54418 standard; DNA; 3046 BP.

XX AAH54418;

XX 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3782.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;

XX endocarditis; ds.

XX Staphylococcus epidermidis.

XX WO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US030782.

XX 09-NOV-1999; 99US-0164258P.

XX (GLAX) GLAXO GROUP LTD.

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

XX useful for vaccinating against infections, e.g. endocarditis.

PT

```
XX PS Claim 8; Page 1413-1414; 2189pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
XX CC and (II) can have antibacterial activity and therefore can be used in
XX CC vaccination. The nucleic acids (I) may be used to produce the S.
XX CC epidermidis polypeptides (II) via the production of vectors containing
XX CC them which are used to produce host cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.
XX CC The polypeptides may also be used to assay for other inhibitors of their
XX CC activity and therefore identify compounds that may be used for the
XX CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
XX CC represent oligonucleotide sequences and primers which are used in the
XX CC exemplification of the present invention. N.B. The present invention
XX CC specifically claims all the polynucleotide sequences given in the
XX CC sequence listing of the present specification, however the sequence
XX CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
XX CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
XX CC for SEQ ID NO:4455 to 4464
XX SQ Sequence 3046 BP; 1103 A; 468 C; 436 G; 1039 T; 0 U; 0 Other;
XX Query Match 10.9%; Score 109.4; DB 4; Length 3046;
XX Best Local Similarity 95.0%; Pred. No. 7.7e-11;
XX Matches 113; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGAGTGGAAATCGTGAAGAGTTTAAATTTTAAAGTGAATGGCTTATGTTGTTTAA 60
DB 2705 CGAGTGGAAATCGTGAAGAGTTTAAATTTTAAAGTGAATGGCTTATGTTGTTTAA 2764
QY 61 ATTGCAACTGCGAGCATGTGGAATTAATAGTTCAGTAACTCAAGTAAAGAGTCAATAA 119
DB 2765 ATTGCAACTGCGAGCATGTGGAATTAATAGTTCAGTAACTCAAGTAAAGTCAATAA 2823

RESULT 36
ACA32000
ID ACA32000 standard; DNA; 900 BP.
XX AC ACA32000;
XX AC ACA32000;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #13657.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Enterobacter cloacae.
XX PN WO200271783-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX PI WPI; 2003-029926/02.
XX PR P-PSDB; ABU28130.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 19870; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway;
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification. The present sequence is
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 900 BP; 198 A; 266 C; 256 G; 180 T; 0 U; 0 Other;
XX Query Match 10.2%; Score 103.2; DB 8; Length 900;
XX Best Local Similarity 52.9%; Pred. No. 9.7e-10;
XX Matches 222; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
QY 137 AGCAGGAAGAGGTACTACGAAGTACCTAAACACCCCTAAACCGTGTGTTCTTGAAGT 196
DB 74 AGGACGAACACGGCAGCGTTTACACTCGATATAAACGCCACACCGATTGCTGCTGGAAC 133
QY 197 ATTCATTTGTTGATCGGTAGTTCGTTTAGATGTTAAACCTGTTGGGATAGCGATGATA 256
DB 134 TCTCGTTGCGCGATGCGCTGCGCGCGGACGTCGACCGCATCGGTATTGCCGACGATA 193
QY 257 ACAAAAAAATCGTATTATTAAACCATTAAGAGATATAAATTTGGAATAATACACTTCTGTAG 316
DB 194 ACGATGCAAAACGGCATCTGCGCCGACGTCGTCGGCAGCTGAACCGTGGCAGTCCGTG 253
QY 317 GAACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAACTTAAACCGAGATTATTTATG 376
DB 254 GAACGCGCGCGCAGCGGAGCTTGAAGCCATTGCTGCCCTGAAACACGACCTGATCATTTG 313
QY 377 CTGATAATATAGACACAAAGGATTATTATAAGACTTTAAATAAATTTCTCTACGATTG 436
DB 314 CCGACAGCAGTCGCCATGCGGGGATTTACACTGCTTGCAGCAAAATCGCGCGGTACTGCG 373
QY 437 AACTGAAAAGTTTCGATGGAGATTATATGAATAATTAATGATGCTTTTAAAAACAATTTCAA 496
DB 374 TGCTTAAGTCCGCAACGAACACCTACGCTGAAATTTGCACTCAGCGGCTATCATCGCGG 433
QY 497 AAGCTTTTAGGTAAAGAAAGAGAGGTAAAAACCGTTAGAGAAACACGATTAAGAAAAATTG 556
DB 434 AAGTGGTAGGTAAAAAACGAGAGATGTCAGGCGCGCTCTGGAACCAACATTAAGAGAGATGG 493
```

```

RESULT 37
ACF70890
ID ACF70890 standard; DNA; 909 BP.
AC ACF70890;
XX
DT 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens nucleotide sequence #9357.
XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough; gene; ds.
XX
OS Photorhabdus luminescens.
XX
PN WO200294867-A2.
XX
PD 28-NOV-2002.
XX
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
PA (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A,
PI Buchrieser C;
XX
PI WPI; 2003-148459/14.
XX
DR
XX
PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
PS Claim 2; SEQ ID NO 9357; 1205pp; French.
XX
CC The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 909 BP; 230 A; 258 C; 225 G; 196 T; 0 U; 0 Other;
Query Match 9.6%; Score 96.8; DB 10; Length 909;
Best Local Similarity 51.6%; Pred. NO. 1.4e-08;
Matches 221; Conservative 0; Mismatches 207; Indels 0; Gaps 0;
QY 137 AGCACGAGAGCTACTACGAAAGTACCTAAACACCTTAACCTGTGTGTCTTGAGT 196
DB 80 AGGATGACGAGGCGAGTTTACCCCTCAATACCGTTCCCGACGCGTAGTGTGCTGGAAT 139
QY 197 ATTCATTGTGTGCGTTAGTTGCTTTAGATGTTAAACCTGTGGGATAGCGGATGATA 256
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
140 TGTCAATTTGCCGATGCATTGGCCGCCATCAATATACAGTCGCGTAGGTATTGCCGATGATA 199
QY ACACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTCGAAAAATACACTTCTGTAG 316
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 ACGATCCCGAGGCGATTCTCACCAGTGTTCGCCAGCGTATTAAAGCCCTGGCAATCAACG 259
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACACAGATTAAATTATG 376
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 GCACCTCGCGCGCAACCAAGCGCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATG 319
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 377 CTGATAATAATAGACACAAAGGTATTATAAGACTTAAATAAATAATGCTCTACGATTG 436
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320 CAGACAGTCAACGCCACCGCGGATTTATCGGCGCACTCAAGGGCATCGCTCTGCTCTGC 379
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
437 AACTGAAAAGTTTCGATGCGAGATTATAATGAAAAATATTGATGCTTTTAAAAACAATTTCAA 496
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 TGCTAAATCCGCAACGAAACCTACGAAAGAGAACTGCAATCTGCCGCGATCATCGGCA 439
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
497 AAGCTTTTAGGTAAGAAAGAGGTAAACCGTTAGAGAAACACGATAGAAAAATTG 556
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 AAGTCATGGGAAAGAGAGTGAATGCAGAAACGCGTGCAGAGACCCGCGAACCATGA 499
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 557 AAGATAT 564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 AAGCTAT 507
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 38
ADR20887/c
ID ADR20887 standard; DNA; 25860 BP.
XX
AC ADR20887;
XX
DT 07-OCT-2004 (first entry)
XX
DE Photorhabdus luminescens genomic DNA encoding two insecticide proteins.
XX
KW insecticide protein; transgenic plant; insect-resistant plant; wheat;
KW maize; gene; ds.
XX
OS Photorhabdus luminescens subsp. laumondii.
XX
FH Key Location/Qualifiers
FT CDS 20872..21306
FT /tag= a
FT /product= "Photorhabdus luminescens insecticide protein
FT #1"
FT CDS 21345..22598
FT /tag= b
FT /product= "Photorhabdus luminescens insecticide protein
FT #2"
XX
WO2003087377-A1.
XX
23-OCT-2003.
XX
17-APR-2003; 2003WO-FR001239.
XX
17-APR-2002; 2002FR-00004798.
XX
(INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
WPI; 2003-833739/77.
XX
P-PSDB; ADR20889, ADR20891.
XX
PT New nucleic acid from Photorhabdus luminescens, useful for producing
PT insecticidal polypeptides active against Lepidoptera, and for producing
PT insect resistant transgenic plants.
XX
PS Claim 22; SEQ ID NO 1; 44pp; French.
XX

```


The invention comprises the amino acid and coding sequences of proteins from *Photobacterium luminescens* which are active against insects. The proteins of the invention are toxic to Lepidoptera, specifically the genera *Plutella*, *Heliothis*, *Helicoverpa*, *Spodoptera* and *Ostrinia*. The DNA and protein sequences of the invention are useful in the preparation of transgenic, insect-resistant plants, specifically wheat and maize. The present *Photobacterium luminescens* genomic DNA contains the coding sequence for the two insecticide proteins of the invention.

Sequence 25860 BP: 7315 A; 5392 C; 6047 G; 7106 T; 0 U; 0 Other;

Query Match 9.6%; Score 96.8; DB 11; Length 25860;
Best Local Similarity 51.6%; Pred. No. 1.8e-08;
Matches 221; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

Qy 137 AGCACGAAGAGTACTACGAAGTACTAAACACCTAAACGTGTTGTTGTTCTTCAGT 196

Dh 8497 AGGATGACAGGCAGTTTACCTCAATAACGTTCCACGCGCTAGTCGTGCTGGAAT 8438

QY 197 ATTCAATTTGGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGATAGCGGATGATA 256

9437 TGTCTATTTGGCGGATGATTGGCGCCATCAATATCACTCCGGTAGGTATTGCGGATGATA 8378

Qy 257 AAAAAAAAAATCGTATTATTATTAACCATTAAGAGATAAAATGGAATAACACTCTCTAG 316

QY 317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAACTTTAAACCCAGATTTAATTG 376

QY 377 CTGATAATAATAGACACAAAGGATTTTATAAGACTTAAATAAAATGCTCTCTACGATTG 436

437 AACGTGAAAAGTTTCGATCGAGATATTAATGAAAATATTGATGCTTTTAAAAACAATTTCAA 496

QY 497 AAGCTTTTAGGTTAAGAGGAAGGTTAAAAACGCTTTAGAAGAACACGATATAGAAAAATTG 556

QY 557 AAGAATAT 564

RESULT 39

Continuation (40 of 57) of ACF67367 from base 3900001 (Photorhabdus luminescens nucleotide sequence)

WP	Fragment Name	Begin	End
1	Fragment 1	100	200
2	Fragment 2	200	300
3	Fragment 3	300	400
4	Fragment 4	400	500
5	Fragment 5	500	600
6	Fragment 6	600	700
7	Fragment 7	700	800
8	Fragment 8	800	900
9	Fragment 9	900	1000

WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000

WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5710000
WP	ACF67367_57	5700001	5810000
WP	ACF67367_58	5800001	5910000
WP	ACF67367_59	5900001	6010000
WP	ACF67367_60	6000001	6110000
WP	ACF67367_61	6100001	6210000
WP	ACF67367_62	6200001	6310000
WP	ACF67367_63	6300001	6410000
WP	ACF67367_64	6400001	6510000
WP	ACF67367_65	6500001	6610000
WP	ACF67367_66	6600001	6710000
WP	ACF67367_67	6700001	6810000
WP	ACF67367_68	6800001	6910000
WP	ACF67367_69	6900001	7010000
WP	ACF67367_70	7000001	7110000
WP	ACF67367_71	7100001	7210000
WP	ACF67367_72	7200001	7310000
WP	ACF67367_73	7300001	7410000
WP	ACF67367_74	7400001	7510000
WP	ACF67367_75	7500001	7610000
WP	ACF67367_76	7600001	7710000
WP	ACF67367_77	7700001	7810000
WP	ACF67367_78	7800001	7910000
WP	ACF67367_79	7900001	8010000
WP	ACF67367_80	8000001	8110000
WP	ACF67367_81	8100001	8210000
WP	ACF67367_82	8200001	8310000
WP	ACF67367_83	8300001	8410000
WP	ACF67367_84	8400001	8510000
WP	ACF67367_85	8500001	8610000
WP	ACF67367_86	8600001	8710000
WP	ACF67367_87	8700001	8810000
WP	ACF67367_88	8800001	8910000
WP	ACF67367_89	8900001	9010000
WP	ACF67367_90	9000001	9110000
WP	ACF67367_91	9100001	9210000
WP	ACF67367_92	9200001	9310000
WP	ACF67367_93	9300001	9410000
WP	ACF67367_94	9400001	9510000
WP	ACF67367_95	9500001	9610000
WP	ACF67367_96	9600001	9710000
WP	ACF67367_97	9700001	9810000
WP	ACF67367_98	9800001	9910000
WP	ACF67367_99	9900001	1000000

Query Match 9.6%; Score 96.8; DB 10; Length 110000;

Best Local Similarity 51.6%; Pred. No. 1.9e-08;
Matches 221: Conservative 0; Mismatches 207;
Indels 0; Gaps 0;

QY 137 AGCAGGAAGAGGTACTACGAAAGTACTTAACACCCCTAAACGTTGTTGTTCTTCAGT 196

95107 AGCTATGACCGGACATTTTATCCCTCAATACCGTTCCCGACGCGTAGTTCGTGTCGAAT 9504

Qy	197	ATTCA	TTG	TGAT	CGT	TAG	TGCT	TAA	CCT	GTT	GGC	GAT	CGG	ATG	ATA	256
Db	95047	TGT	CAT	TGG	CCG	ATC	CAAT	CAAT	CAG	TCG	TAG	GAT	TGG	CGG	ATG	9498

QY	257	ACAAAAAATCGTATTATTAAACCAATTAAGAGATATAATCGAAATACACTTCTGTAG	316
Db	94987	ACGATCCGACGCCATCTCAGATGTTCCGCGGGTATTAAAGCCCTGGGAATCAACCG	94992

QY 317 GAACACGTGAAGCAACCTTAACTTGAAGAAATCAGATAA CTTAAACCGAGATTTAAATTAATTG 316

Db 94927 GCACCTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATTTG 94866

[illegible]

94807 TGCTAAATCCCGAAACGAAACCTACGAAGAGAACTCTCAATGCGCGATCATCGCA 9474

DB	94747	ARGTCATGGGAAAGAGAGT	GAGATGCAGAAACGGCTGGCAGAGCACCGTGAACGATGA	9468
OV	557	AAGAAATAT	564	

Db 94687 AGGCTAT 94680


```
RESULT 40
ACF65388_08
Continuation (9 of 13) of ACF65388 from base 800001 (Photorhabdus luminescens nucleotide
WP Sequence split into 13 fragments LOCUS ACF65388 Accession ACF65388
WP Fragment Name Begin End
WP ACF65388_00 1 110000
WP ACF65388_01 100001 210000
WP ACF65388_02 200001 310000
WP ACF65388_03 300001 410000
WP ACF65388_04 400001 510000
WP ACF65388_05 500001 610000
WP ACF65388_06 600001 710000
WP ACF65388_07 700001 810000
WP ACF65388_08 800001 910000
WP ACF65388_09 900001 1010000
WP ACF65388_10 1000001 1110000
WP ACF65388_11 1100001 1210000
WP ACF65388_12 1200001 1225559

Query Match 9.6%; Score 96.8; DB 10; Length 110000;
Best Local Similarity 51.6%; Pred. No. 1.9e-08; Mismatches 207; Indels 0; Gaps 0;
Matches 221; Conservative 0;

QY 137 AGCACGAAGAAGTACTACGAAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGAGT 196
DB 29830 AGGATGAGCAGCGCAGTTTACCTCAATACCGTTCCCGCAGCGGTAGTCGTCTGGAAT 29889

QY 197 ATTCAATTTCTGATCGTTAGTTAGTTGTTTAACTGTTGGGATAGCGGATGATA 256
DB 29890 TGTCAATTTCCGATGCAATGGCGCCATCAATATCAGTCGGTAGGTATTGCCGATGATA 29949

QY 257 ACAGAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAAATACACTTCTGTAG 316
DB 29950 ACGATCCGCGCGCATTTCTCAGCGATGTTCCGCGGTATTAAAGCCCTGGCAATCAACCG 30009

QY 317 GAACACGTAAGCAACCTAACTTAGAGAAATCAGTAAACTTAAACCCAGATTTAAATTATTG 376
DB 30010 GCATCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGTGAACCTGATCTCATCATTTG 30069

QY 377 CTGATAAATAAGACACAAGGTATTATTAAGACTTAAATAAAATTTGCTCTCAGATTG 436
DB 30070 CAGACAGTCAACGCCACGCGGGATTTATCGGCACTCAAGGGCATCGCTCTCTCTGCTGC 30129

QY 437 AACTGAAAAGTTTCGATGAGATTATAATGAAAAATTGATGCTTTTAAAACAATTTCAA 496
DB 30130 TGTCTAAATCCCGCAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCGATCATCGGCA 30189

QY 497 AAGCTTTAGTAAAGAGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAA 556
DB 30190 AAGTCATGGGGAAGAGAGTGAATGTCAGAAACCGCTGGCAGAGACCGTGAACGATGA 30249

QY 557 AAGAATAT 564
DB 30250 AAGGCTAT 30257
```

Search completed: August 26, 2005, 04:36:53
Job time : 653 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 03:38:21 ; Search time 3906 Seconds
(without alignments)
9823.037 Million cell updates/sec

Title: US-10-724-972A-2580

Perfect score: 1008

Sequence: 1 ggaggggaatcagtgagagg.....gtaaaaagataataagtaa 1008

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g881: *
9: gb_g882: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96.8	9.6	801	9	CL674627
2	95.2	9.4	625	7	CF307287 HDAL-06-
3	87.4	8.7	1566	9	CG757757 P053-1-D0
4	87.4	8.7	1632	9	CL082569 CH216-167
5	86.6	8.6	1135	9	CNS033GQ
6	86.4	8.6	1101	9	CNS0021J
7	83.6	8.3	773	8	BH651171 BOMJP20TR
8	83.6	8.3	1172	3	CR734084 Tetraodon
9	82.4	8.2	1362	9	AG429915 Mus muscu
10	81.6	8.1	1042	9	CL461494 SAIL_1148
11	81.4	8.1	1017	9	CL101048 ISB1-360A
12	81.4	8.1	1626	7	CF238805 AGENCOURT
13	81.2	8.1	1594	9	CL038406 CH216-46A
14	80.8	8.0	1101	9	AL073856 Drosophil
15	80.8	8.0	1519	9	AG386893 Mus muscu
16	80.6	8.0	1402	9	CL078552 CH216-151
17	80.4	8.0	1506	9	AG278469 Mus muscu
18	80.2	8.0	1225	9	CG745927 P039-1-A0
19	80	7.9	1256	9	CL119201 ISB1-76J1
20	79.6	7.9	1231	9	CL694436 PRI0164B
21	79.6	7.9	1231	9	CG755650 P051-1-G0
22	79.6	7.9	1251	9	AG332167 Mus muscu
23	79.2	7.9	625	9	AL229763 Tetraodon
24	79.2	7.9	1101	9	AL108642 Drosophil

4

C 25	79.2	7.9	1353	9	CG744812	CG744812	P037-3-B0
C 26	79	7.8	1392	9	CG757503	CG757503	P052-4-C0
C 27	79	7.8	1536	9	CL078538	CH216-151	CL078538
C 28	78.8	7.8	1162	9	CL077122	CH216-143	BF273407
C 29	78.8	7.8	1185	2	BF273407	GA_Eb001	CL078190
C 30	78.8	7.8	1217	9	CL078190	CH216-148	CG754010
C 31	78.8	7.8	1276	9	CG754010	P049-2-C0	AG350118
C 32	78.8	7.8	1472	9	AG350118	Mus muscu	CC268594
C 33	78.6	7.8	1352	8	CC268594	CH261-67F	CG753732
C 34	78.6	7.8	1811	9	CG753732	P048-4-G0	CL081338
C 35	78.4	7.8	1163	9	CL081338	CH216-161	CG758143
C 36	78.4	7.8	1297	9	CG758143	P053-3-B1	CG744815
C 37	78.4	7.8	1380	9	CG744815	P037-3-B0	CL110553
C 38	78.4	7.8	1594	9	CL110553	ISB1-53P2	CD388253
C 39	78.2	7.8	971	6	CD388253	AGENCOURT	CL037168
C 40	78.2	7.8	1110	9	CL037168	CH216-42N	AG288477
C 41	78.2	7.8	1277	9	AG288477	Mus muscu	CG750869
C 42	78.2	7.8	1373	9	CG750869	P045-2-E1	AJ591978
C 43	78.2	7.8	1453	9	AJ591978	Arabidops	CV481731
C 44	78	7.7	722	7	CV481731	AGENCOURT	AL441457
C 45	78	7.7	964	9	CNS0758R	T7 end of	

ALIGNMENTS

RESULT 1
CL674627
LOCUS
DEFINITION
CL674627 801 bp DNA linear GSS 09-JUL-2004
PRI0112c_G03_2 - PRI0112c.BR (801) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

ACCESSION
CL674627

VERSION
GSS.

KEYWORDS
Pristionchus pacificus

SOURCE
Pristionchus pacificus

ORGANISM
Pristionchus pacificus

REFERENCE
1 (bases 1 to 801)

AUTHORS
Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.

TITLE
AppADB: an Acadb database for the nematode satellite organism

JOURNAL
Pristionchus pacificus

COMMENT
Nucleic Acids Res. 32 (1), D421-D422 (2004)

CONTACT: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

Location/Qualifiers

1..801

/organism="Pristionchus pacificus"

/mol_type="Genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone_lib="Mixed stage fosmid library of P. pacificus

var. California"

/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 9.6%; Score 96.8; DB 9; Length 801;
Best Local Similarity 51.9%; Pred. No. 3.3e-09;
Matches 218; Conservative 0; Mismatches 202; Indels 0; Gaps 0;
QY 137 AGCACGAAGAGTACTACGAAGTACCTAAACACCTTAACCTGTGTGTTCAGT 196
DB 83 AGGACGAACACGCGACGTTTACACTCGAAACACGACCAACGAGTTGTGTGTGGAAC 142

```

Qy 197 ATTCATTGTTGATCGCTTAGTTCCTTTAGATGTTAAACCTGTTGGGATAGCGATGATA 256
Db 143 TCTCGTTGCGCGATCGCTGCGCGCGTGGGACGTCAGCCCGATCGGTATTGCCGACGATA 202
Qy 257 ACAAATAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAAAAATACACTTCTGTAG 316
Db 203 ACGATGCAAAACGATCTGCGCCGAGTGCGTGGCACCTGAAACCGTGGCAGTCCGTG 262
Qy 317 GAACACGCTAAGCAACCTTAATTAGAGAAATCAGTAAACTTTAAACCGAGATTAAATATTG 376
Db 263 GAACGCGCGCGACGCGAGCCTGGAAGCATTGCGCGCTCTGAAACCGACCTGATCATTTG 322
Qy 377 CTGATAATAATAGACACAAAGGATTTTATAAGACTTAAATAAATTTCTCTAGGATTG 436
Db 323 CCGACAGCAGTCGCCATGCGGGGTTTACATCGCCTTCGACAAATCGCGCGTACTGCG 382
Qy 437 AACTGAAAGTTTCGATGAGATTATATGAATATAATATTGCTTTTAAACAAATTTCAA 496
Db 383 TGCTTAAGTCCGCAACGAAACCTACGCTGAAATTTGCAATCTGCGGCTATCATTCGCG 442
Qy 497 AAGCTTTAGGTAAAGAAAGAGAGTAAAAACGCTTAGAAGAACACGATAGAAAAATTG 556
Db 443 AAATGGTGGGTAAAAAGCGAGAGATGACGACGCGCTGGAACAAACATAAGAGAGGATGG 502

RESULT 2
CF307287/c 625 bp mRNA linear EST 15-AUG-2003
LOCUS HDAL1--06-E11.g1 OshDAC1-overexpressing transgenic rice lambda phage 625
DEFINITION cDNA library 1 (HDAL1) Oryza sativa (japonica cultivar-group) cDNA
clone HDAL1--06-E11, mRNA sequence.
CF307287
CF307287.1 GI:33679048
ACCESSION Oryza sativa (japonica cultivar-group)
VERSION Oryza sativa (japonica cultivar-group)
KEYWORDS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SOURCE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
ORGANISM Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 625)
AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
of BioScience and Genetics Institute, GreenGene Biotech Inc.; Division
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhna@gbio.com, bhna@bio.myongji.ac.kr.

FEATURES
Location/Qualifiers
source
1. 625
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="HDAL1--06-E11"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/lab_host="E.coli SOLR"
/clone_lib="OshDAC1-overexpressing transgenic rice lambda
phage cDNA library 1 (HDAL1)"
/notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; Callus was treated with ABA(20um) for 1hour. cDNA
was inserted into lambda Uni-ZAP XR vector at 5' end with
EcoRI and 3' end with XhoI site. mRNA was derived from
rice Histone Deacetylase overexpression line."

ORIGIN
Query Match 9.4%; Score 95.2; DB 7; Length 625;
Best Local Similarity 51.7%; Pred. No. 6.9e-09;

```

```

Matches 217; Conservative 0; Mismatches 203; Indels 0; Gaps 0;
Qy 137 AGCAGGAAGAGGTACTACGAAGTACTTAAACACCTTAAACCGTGTGTGTTCTTGAGT 196
Db 615 AGGACGAACACGCGACCGTTTACACTCGAAAAAACGCCACACCGATTTGTGTGCTGGAAC 556
Qy 197 ATTCAATTTGTTGATGCGCTTAGTTCCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
Db 555 TCTGTTGCGCGATGCGCTGCGCGCGTGGACGTCATCCGATCGGTATTGCCGACGATA 496
Qy 257 ACAAATAAAATCGTATTATTAAACCATTTAAGAGATAAAATTTGGAAAAATACACTTCTGTAG 316
Db 495 ACGATGCAAAACGATCTGCGCCGAGTGCGTGGCACCTGAAACCGTGGCAGTCCGTGCG 436
Qy 317 GAACACGCTAAGCAACCTTAATTAGAGAAATCAGTAAACTTTAAACCGAGATTAAATATTG 376
Db 435 GAACGCGCGCGACCGCGAGCCTGGAAGCATTGCGCCTCTGAAACCGACCTGATCATTTG 376
Qy 377 CTGATAATAATAGACACAAAGGATTTTATAAGACTTAAATAAATTTCTCTAGGATTG 436
Db 375 CCGACAGCAGTCGCCATGCGGGGTTTACATCGCCTTCGACCAATCGCGCGGTACTGCG 316
Qy 437 AACTGAAAGTTTCGATGAGATTATATGAATATAATTTGATGCTTTTAAAAACAATTTCAA 496
Db 315 TGCTTAAGTCCGCAACGAAACCTACGCTGAAATTTGCAATCTGCGGCTATCATTCGCGCG 256
Qy 497 AAGCTTTAGGTAAAGAAAGAGAGTAAAAACGCTTAGAAGAACACGATAGAAAAATTG 556
Db 255 AAATGGTGGGTAAAAAGCGAGAGATGACGACGCGCTGGAACAAACATAAGAGAGGATGG 196

RESULT 3
CG757757/c 1566 bp DNA linear GSS 24-OCT-2003
LOCUS P053-1-D07.zc Ppa Ecori BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG757757
VERSION CG757757.1 GI:37986636
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
AUTHORS Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M. and Sommer, R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES
Location/Qualifiers
source
1. 1566
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa Ecori BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 8.7%; Score 87.4; DB 9; Length 1566;
Best Local Similarity 38.3%; Pred. No. 2.6e-07;

```


Db 1271 AAAAAATTAATNAABAAAAATATATAAAAAAATAAAAAAATAAAAAAATAAANATTA 1330

Qy 974 TATCTAGAGAGATAGTATAAAAAGATAATAGTAA 1008

Db 1331 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1365

RESULT 5

CNS033GQ 1135 bp DNA linear GSS 01-SEP-2000

LOCUS Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone

DEFINITION 208P24 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL226115

VERSION AL226115.1 GI:7885026

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1

AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.

TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE 20296633

PUBMED 10835645

REFERENCE 2

AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL Genome Res. 10 (7), 939-949 (2000)

MEDLINE 20359837

PUBMED 10899143

REFERENCE 3

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.

FEATURES

source

Location/Qualifiers

1..1135

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="208P24"

/clone_lib="G"

/note="Genoscope sequence ID : COAG208DH12SP1-end : PUC-ORI"

ORIGIN

Query Match 8.6%; Score 86.6; DB 9; Length 1135;

Best Local Similarity 46.0%; Pred. No. 3.8e-07;

Matches 292; Conservative 9; Mismatches 330; Indels 4; Gaps 1;

Qy 254 ATACAAAAAATCGTATTATTAACCATTAAGAGATAAAATTCGAAATACACTTCGTG 313

Db 321 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 380

Qy 314 TAGGAACACGTAGAACCTACTTTAGAGAAATCAGTAACTTAAACCAGATTTATTA 373

Db 381 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 440

Qy 374 TTGCTGATATAATAGACACAAAGGTTATTATATAAGACTTTAAATAAAATTTCTCTACGA 433

Db 441 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 500

Qy 434 TTGAACCTGAAAGTTTCGATGGAGATTATATATGAAAAATATTGATGCTTTTAAAAACAATTT 493

Db 501 TTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAT 560

Qy 494 CAAAAGCTTTAGGTAAAGAGAAGAGTAAAGAAACGCTTAGAAGACACACATTAAGAAA 553

Db 561 TAAAAA-----TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 616

Qy 554 TTGAAGATATATAAAAAAGAAATACTATGGTAAATAAAATCAAAAGGTATTGCTGCAGTAG 613

Db 617 TTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAATTTTA 676

Qy 614 CTGCTAAATCAGGTTTCTGCTCATCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673

Db 677 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 736

Qy 674 AACTAGGTTTAAAGAGCATTAAAGTATGATGTTCTAAAGGTTTAAAGTAAAGTATCTTA 733

Db 737 AATTACATNTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAT 796

Qy 734 AAGGACCTTACTTACAATGAACACTGAACTTTATCTCAAGTGAATCTCTGAGCGTATGT 793

Db 797 AAMAYATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 856

Qy 794 TCATAATGACAAACAAAGCAAGTTCTAACGAACTTCTCACTAAAAAGAACTAGAAAAAGATC 853

Db 857 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAC 916

Qy 854 CTGTATGGAAGAAATTAACCGCTGTGAAAAATCAA 888

Db 917 AAAAAATAAATAAATCTCTAAACCAAAAAAATAAAAAA 951

RESULT 6

CNS0021J

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..1135

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="208P24"

/clone_lib="G"

/note="Genoscope sequence ID : COAG208DH12SP1-end : PUC-ORI"

ORIGIN

Query Match

Best Local Similarity

Matches

Qy

Db

Qy

Db

SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1594)
AUTHORS Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTGAGTGACACTATAG
Class: BAC ends
High quality sequence start: 471
High quality sequence stop: 522.
Location/Qualifiers
1..1594
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-46A6"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

FEATURES
source
CNS00HX9
LOCUS 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR35M04 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL073856
VERSION AL073856
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35M04"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
Query Match 8.0%; Score 80.8; DB 9; Length 1101;
Best Local Similarity 44.7%; Pred. No. 5.5e-06;
Matches 169; Conservative 31; Mismatches 178; Indels 0; Gaps 0;
ORIGIN
232 AACCTGTTGGGATGCGGATGATACCAAAAAAATTCGTATTATTAAACCATTAAGAT 291
449 AAACCTGAAATGTAGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 390
292 AAAATTGGAATATACACTTCTGTAGAACACGTAGCACTTAACCTTAGAAGAAATCAGT 351
389 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 330

Query Match 8.1%; Score 81.2; DB 9; Length 1594;
Best Local Similarity 44.0%; Pred. No. 4.6e-06;
Matches 332; Conservative 0; Mismatches 423; Indels 0; Gaps 0;
ORIGIN
254 ATACAAAAAATCGTATTATTAAACCATTAAGACATAAATGGAATAATCACCTTCG 313
756 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 815
314 TAGGAACACGTAGCACTTAACCTAGAGAAATCACTTAACCTTAACACAGATTTA 373
816 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 875
374 TTGCTGATATATAGACACAAGGTATTATTAAGACATTAATAAAATGCTCTACGA 433
876 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 935
434 TTGACTGAAAGCTTTCGATGGAGATTATAATGAAAAATATTGATGCTTTTAAACCAATTT 493
936 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 995
494 CAAAGCTTTAGGTAAGAGAGAGAGTAAAAAAGCGCTTAGAAGAACACGATAAGAAA 553
996 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1055
554 TTGAAGAATATAAAAAAATAAACTATGATATAAAATCAAAAGGTATTGCTCGCATAG 613
1056 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1115
614 CTGCTAAATCAGGTTTGGCTTGCATCCAGCACTCTTATGTTGGTCAATTCCTAAGTC 673
1116 AAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1175
674 AACTAGGTTTAAAGAGCAATTAAGTCATGATGTTACTAAAGCTTTTAAGTAAGTACTCTTA 733
1176 AAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1235
734 AAGGACCTTACTTACAAATGAACACTTTATCTCAAGTGAATCTCTGACGCTATGT 793
1236 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1295


```

Db      455 AAAATATATAAAAAAAAAATATATGTTGTTGTTTATTAATAAAAAAAAAATTTAAA 396
Qy      731 TTAAGGACCTTACTTACAAATGACACTGAAACTTTTATCTCAAGTGAATCCTGAGCGTA 790
Db      395 AAAAAAAAAAAAAAAAAATATAAAAAAAAAAAAAAAAAAGAAATAAAT-AAAAAAAAATAAA 337
Qy      791 TGTTCATATGACAAACAAAGCAAGTCTTACGAACTTCACTAAAGAACTAGAAAAG 850
Db      336 AAAAAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACATACAAAAAAC 277
Qy      851 ATCTGTTATGCAAGAAATTTAA 872
Db      276 AAAACAATAATAAATAATTTAA 255

RESULT 18
CG745927/c
LOCUS   CG745927      1225 bp      DNA      linear      GSS 24-OCT-2003
DEFINITION   P039-1-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
              genomic survey sequence.
ACCESSION   CG745927
VERSION     CG745927.1 GI:37966853
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
            Pristionchus pacificus
ORGANISM    Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1225)
AUTHORS     Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE       An integrated physical and genetic map of the nematode Pristionchus
            pacificus
JOURNAL     Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE    22835951
PUBMED     12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
            Location/Qualifiers
              1..1225
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Ppa EcoRI BAC Library"
                /note="The library was generated by a partial digest of
                the genomic DNA with EcoRI and cloning into the BAC
                vector."

FEATURES             source
  CDS
    1..1225
      /organism="Pristionchus pacificus"
      /mol_type="genomic DNA"
      /strain="California"
      /db_xref="taxon:54126"
      /clone_lib="Ppa EcoRI BAC Library"
      /note="The library was generated by a partial digest of
      the genomic DNA with EcoRI and cloning into the BAC
      vector."

ORIGIN
Query Match      8.0%; Score 80.2; DB 9; Length 1225;
Best Local Similarity 41.9%; Pred. No. 7.2e-06;
Matches 316; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

Qy      254 ATACAAAAAATCGTATTATTAACCATTAAGAGATAAAATTGAAAAATACACTTCTG 313
Db      1169 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1110
Qy      314 TAGGAACACGTAAGCAACTTAAGTAAAGAAATCAGTAACTTAACACGAGTTTAATTA 373
Db      1109 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1050
Qy      374 TTGCTGATAATAATAGACACAAAGTATTATTAAGAGCTTAATAATAATTTGCTCTACA 433
Db      1049 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 990
Qy      434 TTGAACGAAAAAGTTTCGATGGAGATTATATGAAAAATATTGATGCTTTTAAACAAATTT 493

```

```

Db      989 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 930
Qy      494 CAAAAGCTTTAGGTAAAGAGAGAGAGGTAAAAACGCTTAGAAGAACACAGTAAAGAAA 553
Db      929 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 870
Qy      554 TTGAAGAAATATAAAAGAAATTAACATATGGTAAAAATCAAAAGGTATTGCTCGAGTAG 613
Db      869 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810
Qy      614 CTGCTAAATCAGGTTTCTGCTCATCCAAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
Db      809 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 750
Qy      674 AACTAGGTTTTAAAGAGCAATTAAAGTGATGATGTTTACTAAAGGTTTAAAGTATCTTA 733
Db      749 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 690
Qy      734 AAGGACCTTACTTACAAATGAACACTGAACTTTTATCTCAAGTGAATCCTGAGCGTATGT 793
Db      689 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 630
Qy      794 TCATAATGACAAACAAAGCAAGTTCTAACGCAACCTTCACATAAAAGAACTAGAAAAAGATC 853
Db      629 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 570
Qy      854 CTGTATGGAAGAAATTAACGCTCTGTGAAAAATCAACGTTGTTGATATTTTGAACCGTACT 913
Db      569 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 510
Qy      914 TATGGCAAGATCACGTTGTTTAAATTTCTTCAGAGAAATGCGCAAGAACTTGTGTAAT 973
Db      509 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 450
Qy      974 TATCTAAGAAAGATAGTAAAGAAAGATAATAGTAA 1008
Db      449 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415

RESULT 19
LOCUS   CL119201      1256 bp      DNA      linear      GSS 05-JAN-2004
DEFINITION   ISB1-76J12 Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-76J12,
              genomic survey sequence.
ACCESSION   CL119201
VERSION     CL119201.1 GI:40612836
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
            Xenopus tropicalis
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
            1 (bases 1 to 1256)
            Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@wustl.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 233
            High quality sequence stop: 402.
            Location/Qualifiers
              1..1256
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-76J12"

FEATURES             source
  CDS
    1..1256
      /organism="Xenopus tropicalis"
      /mol_type="genomic DNA"
      /db_xref="taxon:8364"
      /clone="ISB1-76J12"

```


TITLE An integrated physical and genetic map of the nematode *Pristionchus pacificus*
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1. .1231
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC library"
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

FEATURES
source
1. .1231
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-122E04.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 7.9%; Score 79.6; DB 9; Length 1231;
Best Local Similarity 41.1%; Pred. No. 9.6e-06;
Matches 310; Conservative 0; Mismatches 445; Indels 0; Gaps 0;
Qy 254 ATACACAAAAATCGTATTATTAACCATTAAGAGATAAAATTCGAAATACACTTCTG 313
Db 1228 AA 1169
Qy 314 TAGGAACACGTAAGCAACCTCACTTAGAAGAAATCAGTAACTTAAACACAGATTAAATTA 373
Db 1168 AA 1109
Qy 374 TTGCTGTAATAATAGACACAAAGGTATTATATAAGCTTTAAATTAATTTGCTCTCAGCA 433
Db 1108 AA 1049
Qy 434 TTGACTGAAAGTTTCGATCGAGATTATATGAATATTGATGCTTTTAAACCAATTT 493
Db 1048 AA 989
Qy 494 CAAAGCTTTAGGTAAAGAGAAGAGTAAAAACGCTTAGAAGAACACGATTAAGAAAA 553
Db 988 AA 929
Qy 554 TTGAAGATATAAAAAAGAAATACTATGGATAAAAATCAAAAGGTATTGCTGCGAGTAG 613
Db 928 ANAAA 869
Qy 614 CTGCTTAATCAGTTTGTCTGCTCATCAAGCACTCTTATGTTGGTCAATTCCTAAGTC 673
Db 868 AA 809
Qy 674 AACTAGGTTTAAAGAACATTAAAGTGATGTTTACTTAAAGGTTTAAAGTATCTTTA 733
Db 808 NNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 749
Qy 734 AAGGACCTTACTTACAAATGAACACTGAAACCTTTATCTCAAGTGAATCCTGAGCGGTATGT 793
Db 748 AA 689
Qy 794 TCATATGACAAACAAAGCAAGTTCTACGACCTTCTCAAGAACTTAGAAAGATC 853
Db 688 AA 629
Qy 854 CTGTATGGAAGAAATTAACCGTGTGAAAAATCAACGTTGTTGATATTTTAGACCGGTGACT 913
Db 628 NNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 569

Qy 914 TATGGCGAAGATCAGGTGGTTTAAATTTCTTCAGAGAAATGCGCAAGAACTGTGTGAAT 973
Db 568 AA 509
Qy 974 TATCTAAGAAAGATAGTAAAGAAAAAGATAATAGTAA 1008
Db 508 AA 474

RESULT 22
AG332167/c 1251 bp DNA linear GSS 02-JUN-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-122E04.T7, genomic survey
DEFINITION sequence.
ACCESSION AG332167
VERSION AG332167.1 GI:47905477
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1251)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Tsukuba Institute, Bio Resource Center,
Tsukuba, Ibaraki, 305-0074 Japan
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY : pBACe3.6
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .1251
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-122E04.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

FEATURES
source
1. .1251
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-122E04.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 7.9%; Score 79.6; DB 9; Length 1251;
Best Local Similarity 51.0%; Pred. No. 9.6e-06;
Matches 175; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
Qy 254 ATACACAAAAATCGTATTATTAACCATTAAGAGATAAAATTCGAAATACACTTCTG 313
Db 1192 ATTATAAATAAAAAAAAAATAATNAAAAAAAAAATAATAATAATAATAATAATAATA 1133
Qy 314 TAGGACACGTAAGCAACCTCACTTAGAAGAAATCAGTAAACTTAAACACAGATTAAATTA 373
Db 1132 TATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1073
Qy 374 TTGCTGTAATAATAGACACAAAGGTATTTATAAGACTTTAAATAAAATTCCTCTACGA 433
Db 1072 TAAAAAATAATAAAAAATATATTTAAAAATATATATAAAAAATAAAAAATAAAAAATA 1013

[illegible]

RESULT 23
CNS036A2 625 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION T15D15 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL229763
VERSION AL229763.1 GI:7889758
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

```

ORIGIN
          7.9%; Score 79.2; DB 9; Length 625;
Query Match
Best Local Similarity 48.2%; Pred. No. 1.1e-05;
Matches 201; Conservative 6; Mismatches 210; Indels 0; Gaps 0;
QY 180 TGTGTGTTCTTCGAGTATTTCATTTGTTGATCGCTAGTTGCTTTAGATGTTAAACCTGT 239

```

[illegible]

```

ORIGIN                               7.9%; Score 79.2; DB 9; Length 1101;
Query Match                        48.6%; Pred.No. 1.1e-05;
Best Local Similarity              22; Mismatches 160; Indels 2; Gaps 1;
Matches 174; Conservative
Qy      239 TTGGGATAGCGGATGATAACAAAAAATTCGTATTATTAAACCATTAGAGATAAAATTG 298
||| ||| ||| ||| ||| ||| ||| : || | ||| ||| ||| ||| ||| ||| |||

```


Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1..1392
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

LOCUS CL078538 1536 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-151C11_Sp5.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-151C11, genomic survey sequence.
VERSION CL078538.1 GI:40534451
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp5 atctgcgcttgcgaccc
Class: BAC ends
High quality sequence start: 1061
High quality sequence stop: 1145.
Location/Qualifiers
1..1536
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-151C11"
/sex="male"
/cell_line="Stock 248 P7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

FEATURES
source
Query Match 7.8%; Score 79; DB 9; Length 1392;
Best Local Similarity 44.9%; Pred. No. 1.3e-05;
Matches 338; Conservative 0; Mismatches 412; Indels . 3; Gaps 1;
QY 254 ATAACAAAAAATCGTATTATTAACCACTTAAGAGATAAAATTCGAAATACACTTCG 313
DB 928 AAAAAATATAAANAANAATAAATAAATTAATAAATTAATAAATAAATAAATAAATA 869
QY 314 TAGGAACAGTAAGCACTACTTAGAAGAAATCAGTAACTTAACAGAGATTTAATA 373
DB 868 TTATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 809
QY 374 TTGCTGATATAATACACAAAGAGTATTATTAAGAGCTTAATAAATTCGCTCAGCA 433
DB 808 AATAAAAAATAAANAANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 749
QY 434 TTGAACCTGAAGATTTGCGATGAGATTTAATAACAAATATTGATGCTTTTAAACAAATTT 493
DB 748 AATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 689
QY 494 CAAAGCTTTAGGTAAAGAAAGAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 553
DB 688 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 629
QY 554 TTGAAGATATAAAGAAAGAAATCTATGATTAATAAATCAAAAGGTATTGCTGCAGTAG 613
DB 628 TAAATATTTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 569
QY 614 CTGCTAAATCAGGTTTGCTTGCTCAGCAAGCACTCTTATGCTGCTCAATTCCTAAGTC 673
DB 568 AATTTTAAATAAATTTATATTAATTTATTTATTTATTTATTTATTTATTTATTTAT 509
QY 674 AACTAGGTTTTAAAGAGCAATTAAGTGATGATGTTACTAAAGGTTTAAAGTAAAGTCTTA 733
DB 508 TTATATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 449
QY 734 AAGGACCTTACTTACAAATGACACACTGAACTTTATCTCAAGTGATCTCGAGCGTATCT 793
DB 448 AATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 389
QY 794 TCATATATGACAAACAAAGCAAGTCTTAACGCACTTCACTAATAAGCACTAGAAAAAGATC 853
DB 388 TTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 329
QY 854 CTGTATGGA---AGAAATTAACGCTGTGAAATAATCAACGCTGTGATATTTTAAAGCGG 910
DB 328 TAATTTATATTTATTTATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATA 269
QY 911 ACTTATGGCAGATCAGCTGGTTTAAATTTCTTCAGAGAAATGCGAAAGAACTTGTTG 970
DB 268 ATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 209
QY 971 AATTATCTAAGAAAGATAGTAAAAAGATAATA 1003
DB 208 AATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 176

ORIGIN
Query Match 7.8%; Score 79; DB 9; Length 1536;
Best Local Similarity 51.3%; Pred. No. 1.3e-05;
Matches 178; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
QY 254 ATAACAAAAAATCGTATTATTAACCACTTAAGAGATAAAATTCGAAATACACTTCG 313
DB 912 AAAAAATATAAANAANAATAAATAAATTAATAAATAAATAAATAAATAAATAAATAA 971
QY 314 TAGGAACAGTAAGCACTACTTAGAAGAAATCAGTAACTTAACAGAGATTTAATA 373
DB 972 AAAAAATATAAANAANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1031
QY 374 TTGCTGATATAATAGACACAAAGGTATTATTAAGACTTTAATAAATTCGCTCAGCA 433
DB 1032 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1091
QY 434 TTGAACCTGAAGATTTGCGATGAGATTTAATAACAAATATTGATGCTTTTAAACAAATTT 493
DB 1092 ATAAAAAATAAANAANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1151
QY 494 CAAAGCTTTAGGTAAAGAAAGAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 553
DB 1152 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1211
QY 554 TTGAAGATATAAAGAAAGAAATCACTATGATGATAAATAAATAAATAAATAAATAA 600
DB 1212 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1259

RESULT 28
CL077122
LOCUS CL077122 1162 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-143B5_Sp6.1 CH216 Xenopus tropicalis genomic clone
CH216-143B5, genomic survey sequence.

[illegible]

[illegible]

RESULT 30			
CL078190/c			
LOCUS	CL078190	1217 bp	DNA linear
DEFINITION	CH216-148G2_Sp6.1 CH216 <i>Xenopus tropicalis</i> genomic clone		
	CH216-148G2 genomic survey sequence.		

KEYWORDS
GSS.
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoides; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 1217)
AUTHORS
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE
A physical map of the xenopus tropicalis genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Richard K Wilson

FEATURES	Location/Qualifiers
source	1. .1217

ORIGIN	/note="Vector: pTABAC2.1; CHORI-216 Xenopus tropicalis BAC library"	7.8%; Score 78.8; DB 9; Length 1217;	Best Local Similarity 45.0%; Pred. No. 1.4e-05;	Matches 342; Conservative 0; Mismatches 412; Indels 6; Gaps 1;
Qy	251	ATGATTAACAAAAAATCGTATTATTATTAACCATTAAGAGATAAAATGGGAATACACTT	310	
Db	1077	ATAATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	1018	
Qy	311	CTGTAGGAACACGTAGCAACCTTAATTAGAGAAATCAGTAAACTTTAAACCCAGATTAA	370	
Db	1017	ATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	958	
Qy	371	TTATTGCTGATAATAATAGACACAAAGGTATTATTAAAGACTTAATAATAAATTTGCTCTTA	430	
Db	957	AAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	898	
Qy	431	CGATTGAACTGAAAAAGTTTCGATGGAGATTATTAATGAATAATATTGATGCTTTTAAAAACAA	490	
Db	897	AAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	838	
Qy	491	TTTCAAAAGCTTTAGGTAAAGAGAGAAAGGTAAAAAAGCGTTTAGAAGACACGATTAAGA	550	
Db	837	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	778	
Qy	551	AAATTGAAGAAATATAAAAAAGAAATAACTATGATTAATAAATCAAAAGGTATTGCCTCGAG	610	
Db	777	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	718	
Qy	611	TAGCTGCTAAATCAGGTTTCTGCTTCATCCAAAGCAACTCTTATGTGGTCAATTCCTAA	670	
Db	717	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	658	
Qy	671	GTCAACTAGGTTTTTAAAGAGCATTAAGTGATGATGTACTTAAAGGTTTAAAGTATC	730	
Db	657	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAA	598	
Qy	731	TTAAAGGACCTTACTTCAATGACACTGAAACTTTTATCTCAAGTGAATCTCTGAGCGTA	790	
Db	597	AGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	538	
Qy	791	TGTTCTAATAGCAAAACAAAGCAAGTTCTTAACCGAACCTTC-----ACTAAAGAACTAG	844	
Db	537	CGCCAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	478	
Qy	845	AAAAAGATCCTGTATGGAAGAAATTAACCGTGTGAAAAAATCAACGTGTTCGATATTATTAG	904	
Db	477	AAAAAGAAAAAATAAGAACAAAAAATAAAAAAGGGGAGAAAAAATAAAAAAAGGGGTTTTTTT	418	
Qy	905	ACCGTGACTTATGGCGAAGATCAGCTGTTTTTAAATTTCTTCAGAAGAAATGGCAAAAGAAC	964	
Db	417	TTTTTCCAAAAAATAAAAAAATAAGGGGGGGTTTTTTCTTTCTTAAAAAATAAAAAAATGT	358	
Qy	965	TTGTTGAATTATCTTAAGAAAGATAGTAAAAAAGATAATAA 1004		
Db	157	TTTACGACCAATCGAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	318	

RESULT 31	
CG754010/c	
LOCUS	
DEFINITION	CG754010 1276 bp DNA linear GSS 24-OCT-2003 P049-2-C03.ya Ppa ECORI BAC Library Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION	CG754010
VERSION	CG754010
KEYWORDS	CG754010.1 GI:37979071 GSS.
SOURCE	Pristionchus pacificus
ORGANISM	Pristionchus pacificus Eukaryota; Metazoa; Nematoda; Diplogasterida; Necodiplogasteridae; Pristionchus.

LOCUS DEFINITION	CG753732	1811 bp	DNA	linear	GSS 24-OCT-2000
ACCESSION	P048-4-G03.ya Ppa EcoRI BAC Library				
VERSION	CG753732				
KEYWORDS	CG753732.1	GI:37978509			
SOURCE ORGANISM	GSS. Pristionchus pacificus Pristionchus pacificus Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus. 1 (bases 1 to 1811)				
REFERENCE	Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perbolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J. An integrated physical and genetic map of the nematode Pristionchus pacificus				
AUTHORS	Mol. Genet. Genomics 269 (5), 715-722 (2003)				
JOURNAL	22835951				
MEDLINE	12884007				
PUBMED	Contact: Sommer RJ				
COMMENT	Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel: 00497071601371 Fax: 00497071601498 Email: ralf.sommer@uebingen.mpg.de Class: BAC ends.				
FEATURES	Location/Qualifiers. 1..1811 /organism="Pristionchus pacificus" /mol_type="genomic DNA" /strain="California" /db_xref="taxon:54126" /clone_lib="Ppa EcoRI BAC Library" /note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."				
ORIGIN	Query Match 7.8%; Score 78.6; DB 9; Length 1811; Best Local Similarity 37.4%; Pred. No. 1.5e-05; Matches 282; Conservative 0; Mismatches 473; Indels 0; Gaps 0;				
Qy	254 ATACAAAAAATTCGTATTATTATTAACCATTTAAGAGATAAAATTCGAAATATACACTTCTG 313				
Db	1253 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1194				
Qy	314 TAGAACACGTAAACCACTTAAGAGAAATCAGTAACTTAAACCGAGATTAAATTA 373				
Db	1193 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1134				
Qy	374 TTGCTGATATAATAGACACAAAGGTATTATTAAGACTTAAATAAAATTCCTCTACGA 433				
Db	1133 NANNANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1074				
Qy	434 TTGAATCGAAAAAGTTTCGATGGAGATTATTAATGAAAAATATTGATGCTTTAAAAACAATTT 493				
Db	1073 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1014				
Qy	494 CAAAGCTTTAGTAAAGAAAGAGGTAAAAAACGCTTAGAAGAAACACGATAAGAAAA 553				
Db	1013 NNNAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 954				
Qy	554 TTGAAGATATAAAGAAATAACTATGGTTAAATAATCAAAAGGTATTGCTGCTAGTAG 613				
Db	953 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 894				
Qy	614 CTGCTAAATCAGGTTTGTCTGCTCATCCAAGCACTCTTATGTTGGTCAATTCTCCTAAGTC 673				
Db	893 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 834				
Qy	674 AACTAGGTTTTAAAGAGCACTTAAGTGATGATGTTACTTAAAGGTTTAAAGTAAAGTACTCTTA 733				
Db	833 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 774				

Qy	734	AAGCACTTACTTACAAATGAACACGACCTGAAACTTTATCTCAAGTGAATCTCGAGGCTATGT	7393
Db	730	AAAAAANNAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	671
Qy	794	TCATAATGACAAACAAGCAAGTCTTCAACGAACCTTCACCTAAAAGAACTAGAAAAAGATC	853
Db	670	AAAAAANNAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	611
Qy	854	CTGTATGGAAGAAATTAACCGCTGTGAAAAATCA	887
Db	610	AAAAAANNAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	577
RESULT 38			
LOCUS	CL110653/c		
DEFINITION	ISB1-53P23 Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23.	1594 bp	DNA linear GSS 05-JAN-2004
ACCESSION	CL110653		genomic survey sequence.
VERSION	CL110653.1	GI:40604288	
KEYWORDS	GSS.		
SOURCE	Xenopus tropicalis (western clawed frog)		
ORGANISM	Xenopus tropicalis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
JOURNAL	Xenopodinae; Xenopus; Silurana.		
COMMENT	1 (bases 1 to 1594)		
	Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,		
	Mardis, E. and Wilson, R.		
	A physical map of the xenopus tropicalis genome		
	Unpublished (2003)		
	Contact: Richard K Wilson		
	Genome Sequencing Center		
	Washington University School of Medicine		
	Email: submissi@wustl.edu		
	Insert Length: 75000 Std Error: 0.00		
	Seq primer: Sp6 ATTTAGGTGACACTATAG		
	Class: BAC ends		
	High quality sequence start: 390		
	High quality sequence stop: 470.		
FEATURES			
source	1. .1594		
	/organism="Xenopus tropicalis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:8364"		
	/clone="ISB1-53P23"		
	/clone_lib="ISB1"		
	/notes="Vector: pBel0BAC11; ISB-1 Xenopus tropicalis BAC		
	Library Segment 1"		
ORIGIN			
Query Match	7.8%;	Score 78.4;	DB 9;
Best Local Similarity	46.2%;	Pred. No. 1.7e-05;	
Matches	295;	Conservative 0;	Mismatches 341; Indels 2; Gaps 1;
Qy	251	ATGATAACAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAAAAATACACTT	310
Db	1432	ATATAAAAAACAAAAATAAAAAATAAATAAAAAAATAAATAAATAAATAAATAAATAA	1373
Qy	311	CTGTAGAACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAACTTAAACCAGATTAA	370
Db	1372	AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1313
Qy	371	TTATTGCTGATAATAATAGACACAAAGGTATTATTAAAGACTTAAATAAATTTGCTCCTA	430
Db	1312	AAAAATATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1253
Qy	431	CGATTGACTGAAAGCTTCGATGGAGATTATTAATGAATAATTTGATGCTTTTAAACAA	490
Db	1252	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1195
Qy	491	TTTCAAAGCTTTAGGTAAAGAGAGAAAGGTAAAAAACCGCTTAGAAGAACACGATAGA	550

/note="Vector: pTABAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

```
Query Match      7.8%; Score 78.2; DB 9; Length 1110;
Best Local Similarity 48.7%; Pred. No. 1.8e-05;
Matches 203; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

Qy 180 TGGTGTCTTCTTGAGTATTCATTTGGTGGCGGTAGCTTGTAGATGTTAAACCTGT 239
Db 933 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
TTTTTTTTTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTT 874

Qy 240 TGGGATACCGATGATACAAAAAATCGTATTATTAAACCATTTAAGAGATATAATTGG 299
Db 873 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
TTGTTGTGAAGATAGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 814

Qy 300 AAAATACACTTCTGTAGGAACACGTAAGCACTTAACCTTAGAAGAAATCAGTAAACTTAA 359
Db 813 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 754

Qy 360 ACCAGATTAAATTATGCTGATAATAATAGACACAAAGGTATTTATAAGACTTTAAATAA 419
Db 753 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 694

Qy 420 AATTGCTCCTACGATTGAACTGAAAGTTTCGATGGAGATTATATGATAAATATTGATGC 479
Db 693 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 634

Qy 480 TTTTAAAAACAATTTCAAAAGCTTTAGGTAAAGAGAGAGGTAAAAACGCTTAGAAGA 539
Db 633 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
AAAGATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 574

Qy 540 ACACGATAGAAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAATCAAAA 596
Db 573 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 517
```

Search completed: August 26, 2005, 06:59:20
Job time : 3914 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 26, 2005, 05:54:23 ; Search time 2119 Seconds
(without alignments)
3112.522 Million cell updates/sec

Title: US-10-724-972A-2580

Perfect score: 1008

Sequence: 1 ggagtgaatcagtgagagg.....gtaaaaagataaagtaa 1008

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008	100.0	1008	22	US-10-724-972A-2580
2	993	98.5	993	17	Sequence 2580, Ap
3	567	56.2	984	17	Sequence 34954, Ap
4	564.4	56.0	1014	9	Sequence 7644, Ap
5	531.8	52.8	927	9	Sequence 8156, Ap
6	309.8	30.7	525	17	US-09-815-242-8156
7	265.6	26.3	960	9	US-10-282-122A-35107
					Sequence 8024, Ap

ALIGNMENTS

RESULT 1

US-10-724-972A-2580
Sequence 2580, Application US/10724972A

Publication No. US20040147734A1

GENERAL INFORMATION:

APPLICANT: Doucette-Stamm, Lynn

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: PATH03-16

CURRENT APPLICATION NUMBER: US/10/724,972A

CURRENT FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: 09/450,969

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: 09/134,001

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 7544

SEQ ID NO 2580

LENGTH: 1008

TYPE: DNA

ORGANISM: S.epidermidis

US-10-724-972A-2580

Query Match 100.0%; Score 1008; DB 22; Length 1008;
Best Local Similarity 100.0%; Pred. No. 3.1e-162;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 4194, Ap
Sequence 301, App
Sequence 7949, Ap
Sequence 355, App
Sequence 355, App
Sequence 2556, Ap
Sequence 2556, Ap
Sequence 30706, A
Sequence 2066, Ap
Sequence 9415, Ap
Sequence 2107, Ap
Sequence 4605, Ap
Sequence 604, App
Sequence 604, App
Sequence 318, App
Sequence 15, Appl
Sequence 15, Appl
Sequence 238, App
Sequence 238, App
Sequence 10009, A
Sequence 19870, A
Sequence 6372, Ap
Sequence 20621, A
Sequence 14, Appl
Sequence 17083, A
Sequence 32601, A
Sequence 5558, Ap
Sequence 6774, Ap
Sequence 33236, A
Sequence 15706, A
Sequence 120013, A
Sequence 102083, A
Sequence 17217, A
Sequence 383, App
Sequence 383, App
Sequence 375, App
Sequence 1, Appl
Sequence 3400, Ap

```
Qy 1 GGAGTGGAAATCAGTGCAGAGGTTTAAAAATTTTAAAGTGAATGGCTTATGTTGTTTAA 60
Db 1 GGAGTGGAAATCAGTGCAGAGGTTTAAAAATTTTAAAGTGAATGGCTTATGTTGTTTAA 60
Qy 61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
Db 61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
Qy 121 GATGAGGTGAAATCAAGCACGAAGAAGGTACTACGAAGAGTACCTTAAACACCCTAAACGT 180
Db 121 GATGAGGTGAAATCAAGCACGAAGAAGGTACTACGAAGAGTACCTTAAACACCCTAAACGT 180
Qy 181 GTTGTGTTCTTGTAGTATTCATTTGTTGATCGGTTAGTGTGCTTTAGATGTTTAAACCTGTT 240
Db 181 GTTGTGTTCTTGTAGTATTCATTTGTTGATCGGTTAGTGTGCTTTAGATGTTTAAACCTGTT 240
Qy 241 GGGATAGCGGATGATAACAAAAAATCGTATTATTAACACCATTAAGAGATAAAATTTGGA 300
Db 241 GGGATAGCGGATGATAACAAAAAATCGTATTATTAACACCATTAAGAGATAAAATTTGGA 300
Qy 301 AAATACACTTCTGTAGGAACCGTAAGCAACCTTAACTTAGAAGAAATCAGTAAACTTAAA 360
Db 301 AAATACACTTCTGTAGGAACCGTAAGCAACCTTAACTTAGAAGAAATCAGTAAACTTAAA 360
Qy 361 CCAGATTAAATTTGCTGATATAATATAGACACAAAGGTATTTATAAGACTTAAATAAA 420
Db 361 CCAGATTAAATTTGCTGATATAATATAGACACAAAGGTATTTATAAGACTTAAATAAA 420
Qy 421 ATTGCTCTCAGATTGAACTGAAAGTTTCGATGAGATTAATATGAAATATTGATGCT 480
Db 421 ATTGCTCTCAGATTGAACTGAAAGTTTCGATGAGATTAATATGAAATATTGATGCT 480
Qy 481 TTTAAAACAATTTCAAAGCTTTAGTAAAGAAAGAGTAAAGGTTAAAAACGCTTAGAAGAA 540
Db 481 TTTAAAACAATTTCAAAGCTTTAGTAAAGAAAGAGTAAAGGTTAAAAACGCTTAGAAGAA 540
Qy 541 CACGATAAGAAAAATGAAGATATAAAAAAGAAATTAATGATGATGATGATGATGATGAT 600
Db 541 CACGATAAGAAAAATGAAGATATAAAAAAGAAATTAATGATGATGATGATGATGATGAT 600
Qy 601 TTGCTGCTGAGTGTCTGATTAAGTGTGCTGCTCATCAAGCAACTCTTATGTTGGT 660
Db 601 TTGCTGCTGAGTGTCTGATTAAGTGTGCTGCTCATCAAGCAACTCTTATGTTGGT 660
Qy 661 CAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTAAGTGAATGATGATGATGATGATGAT 720
Db 661 CAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTAAGTGAATGATGATGATGATGATGAT 720
Qy 721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTGAAT 780
Db 721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAACTTTATCTCAAGTGAAT 780
Qy 781 CCTGAGCTATGTTCAATGACAAACAAAGCAAGTCTTAACGAACTTCACTTAAAGAA 840
Db 781 CCTGAGCTATGTTCAATGACAAACAAAGCAAGTCTTAACGAACTTCACTTAAAGAA 840
Qy 841 CTAGAAAAGAGTCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGATATT 900
Db 841 CTAGAAAAGAGTCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGATATT 900
Qy 901 TTAGACCGTGAATTAAGGCAAGATCAGTGGTTTAAATTTCTTCAGAGAAATGCGCAAAA 960
Db 901 TTAGACCGTGAATTAAGGCAAGATCAGTGGTTTAAATTTCTTCAGAGAAATGCGCAAAA 960
Qy 961 GAACCTGTTGAATTTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
Db 961 GAACCTGTTGAATTTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
```

RESULT 2

US-10-282-122A-34954

, Sequence 34954, Application US/10282122A

```
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EP/TRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34954
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-10-282-122A-34954
```

```
Query Match 98.5%; Score 993; DB 17; Length 993;
Best Local Similarity 100.0%; Pred. No. 11e-159;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 GTGAGAGGTTTAAAAATTTTAAAGTGAATGGCTTATGTTGTTTAAATGCAACTGCA 72
Db 1 GTGAGAGGTTTAAAAATTTTAAAGTGAATGGCTTATGTTGTTTAAATGCAACTGCA 60
Qy 73 GCATGTGGAATATAGTTCAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 132
Db 61 GCATGTGGAATATAGTTCAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 120
Qy 133 ATCAAGCACGAAGAAGGTACTACGAAAGTACTCAAAAGTACTTAAACCGTGTGTTGTTCTT 192
Db 121 ATCAAGCACGAAGAAGGTACTACGAAAGTACTTAAACCGTGTGTTGTTCTT 180
Qy 193 GAGTATTCAATTTGTGATGCGTTAGTTCGTTTATAGATGTTAAACCTGTTGGATAGCGAT 252
Db 181 GAGTATTCAATTTGTGATGCGTTAGTTCGTTTATAGATGTTAAACCTGTTGGATAGCGAT 240
Qy 253 GATTAACAAAAAATCGTATTATTTAAACCATTAAGAGTAAATTTGGAATACACTTCT 312
Db 241 GATTAACAAAAAATCGTATTATTTAAACCATTAAGAGTAAATTTGGAATACACTTCT 300
Qy 313 GTAGAAACACGTAAGCAACCTTAACTTAAAGAAAAATCAGTAAACTTAAACAGATTAAAT 372
Db 313 GTAGAAACACGTAAGCAACCTTAACTTAAAGAAAAATCAGTAAACTTAAACAGATTAAAT 372
```


Db 662 AACTAGGATTTAAATGCAATTAAGTGACGATGAACAAAAGGTTTAAGTAAATTTGA 721
Qy 734 AAGGACCTTACTTCAAAATGAACACTGAAACCTTTATCTCAAGTGAATCCTGAGCGTATGT 793
Db 722 AAGGACCTTACTTCAAAATGAACACTGAAACCTTTATCTCAAGTGAATCCTGAGCGTATGA 781
Qy 794 TCATATGACAAACAAACCAAGTCTTAACGACCTTCACTAAAGAACTAGAAAAGATC 853
Db 782 TCATATGACAGATCATGCTAAAAAAGATCTCTGCTGAATTCGAAGATTTACAAGAAGATG 841
Qy 854 CTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGCTGTGATATTTAGACCGTGAAT 913
Db 842 CAACATGGAAGAAAGTTGAATGCAGTAAATTAATTCGCGTGAATTTGTTGACCGTGAATG 901
Qy 914 TATGGCAAGATCAGTGGTGTAAATTTCTTCAGAGAAATGGCAAAAGAACTTTGTGAAT 973
Db 902 TTTGGCAAGATCTCGTGGCTTAAATTTCTCTGAAGAAATGGCTAAAGAACTTTGTGAAT 961
Qy 974 TATCTAAGAAAGATGTAAAAA 996
Db 962 TATCAAAAAAGAACAAAAAGTAA 984

RESULT 4

US-09-815-242-8156
; Sequence 8156, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIORITY FILING DATE: 2001-03-21
; PRIORITY APPLICATION NUMBER: 60/191,078
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY APPLICATION NUMBER: 60/206,848
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY APPLICATION NUMBER: 60/207,727
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY APPLICATION NUMBER: 60/242,578
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY APPLICATION NUMBER: 60/253,625
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY APPLICATION NUMBER: 60/257,931
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY APPLICATION NUMBER: 60/269,308
; PRIORITY FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8156
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1014)
US-09-815-242-8156

Query Match 56.0%; Score 564.4; DB 9; Length 1014;
Best Local Similarity 73.1%; Pred. No. 1.1e-86;
Matches 724; Conservative 0; Mismatches 266; Indels 0; Gaps 0;

Qy 7 GAATCAGTGAGAGGTTAAAAATTTTAAGTGTAAATGGCTTATGTTGTTTAAATGCA 66

Db 25 GGAACGATGAGAGGTCTAAAAACCTTTTAGTATATTTGGATTAAATAGTTGCTTATTTTA 84
Qy 67 ACTCAGCATGTGGAATTAATAGTTCAAGTAACTCAAGTAAAGAGTCAATCAAAAGATGGA 126
Db 85 GTTCAGCGTTGTGTAATACGGATAATTCAGAGTAAAGAAATCATCAACTAAAGATACT 144
Qy 127 GTTCAAAATCAAGCAGCAAGAGGTACTACGAAAGTACTAAAAACCCCTAAAAACGTTGTGT 186
Db 145 ATTTCGGTAAAGATGAATAGTTCAGTAAAGATCACTAAAGATGCAAAACGTATCGTT 204
Qy 187 GTTCTTGAGTATTCATTTGTTGATCGTGTAGTTGCTTTTAGATGTAAACCTGTTGGGATA 246
Db 205 GTATTAGAGTACTCATTTGCGAGATGCATTTAGCAGCTTAGACGTTAAACACGTTGGTATT 264
Qy 247 GCGATGATTAACAAAAAATCGTATTATTAACCATTAAGAGATAAAATTTGGAATATAC 306
Db 265 GCTGATGATGTAAGAAAAACGTTATCACTTAAACCCAGTTAGAGAAAAATTTGGGAATTTAT 324
Qy 307 ACTTCTGTAGGAACACGTTAAGCAACCTTAACTTAGAGAAATCACTTAAACCTTAAACCCAGAT 366
Db 325 ACTTCTGTAGGTACACGTTAAACAGCCAACTTAGAGGAATTTAGTAAATTTAAACCCGAT 384
Qy 367 TTAATTTATGCTGATTAATAGACACAAAGGTATTTATAAGACTTTAAATTTAAATTTGCT 426
Db 385 TTAATTTATGCTGATAGCAGTAGACATAAAGGTATTAAATAAGAAATTTAAACCAATTTGCA 444
Qy 427 CCTACGATTAACCTGAAAGTTTCGATGGAGATTTATTAATGAATATTTGATGCTTTTAA 486
Db 445 CCAACATTTATCAATTAAGAGTTTGTAGTGGAGACTACAAACCAAAACATTTAAATTCGTTCAA 504
Qy 487 ACAATTTCAAAAGCTTTTAGGTAAAGAGAAAGCAAGGTAAAAACGCTTAGAGAAACACGAT 546
Db 505 ACAATTTGCTAAGCTTTTAAATTAAGAAAGAAAGAGCGGAAAGCGTCTTGTGAACTGAT 564
Qy 547 AAGAAAATTTGAAGATATATAAAAAAATAACTATGATGAATAAAAAATCAAAAGTATTTGCT 606
Db 565 AAATTTAATCAAAAAAGTATAAAGATGAATTAAGTTTGTATAGAAATCAAAAAAGTGTCTCA 624
Qy 607 GCAGTAGCTGCTAAATCAAGTTTGTCTGCTCATCAAGCAACTCTTATGTTGGTCAATTC 666
Db 625 GCAGTTGTTGCTAAAGCTGTTTATTAGCACATCCAACTTATCATATGTTTGGACAAATTT 684
Qy 667 CTAAGTCAACTAGGTTTAAAGAAAGCAATTAAGTGTATGTTTACTTAAAGGTTTAAAGTAAG 726
Db 685 TTAACGAACTTGGATTTAAAAATGCATTAAGTGTATGTAACAAAGGTTTAAAGTAA 744
Qy 727 TATCTTAAAGGACCTTACTTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCCTGAG 786
Db 745 TACTTGAAGGACCTTACTTACAATTAGATACTGAAACATTTAGCTGACTTAAATCCTGAA 804
Qy 787 CGTATGTTCAATGAACAAACAAAGCAAGTTCTAACGAACTTCACTTAAAGAACTAGAA 846
Db 805 CGCATGATTTATATGACAGATAATGCTAAAAAAGATTTCTGCTGAATTCGAAGATTTACAA 864
Qy 847 AAGATCTCTGTATGGAAGAAATTAACCGCTGTGAAAAATCAACGTTGTGATTTTATAGAC 906
Db 865 GAAGATCCAACTTGGAAAAAGTTGAACCGAGTTTAAAAATATTCGCTGATTTGTTGAC 924
Qy 907 CGTGAATTTATGGCAAGATCACTGCTGTTTAAATTTCTTCAGAGAAATTTGCAAAAGAACTT 966
Db 925 CGTGAATTTGCGCAAGATCTCGTGGCTTAAATTTCTTCTGAAGAAATGCTTAAAGAACTT 984
Qy 967 GTTGAATTTATCTAAGAAAGATAGTAAAAA 996
Db 985 GTTGAATTTATCAAAAAAGAACAAAAAGTAA 1014

RESULT 5

US-09-815-242-4404
; Sequence 4404, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert


```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35107
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35107

Query Match      30.7%; Score 309.8; DB 17; Length 525;
Best Local Similarity 76.3%; Pred. No. 2.1e-43;
Matches 399; Conservative 0; Mismatches 112; Indels 12; Gaps 1;

QY 14 TGAGAGGTTTAAATTTTAAAGTAAATGCGTTATGTTGTTTAAATGCAACTGCAG 73
DB 2 TGAAGGCTTTAAAGTTTCAGGTATAGTCGCAATTAATTTGCAATGTTTGTACTG 61

QY 74 CATGTGGAATAATAGTTCAGTAACTCAAG-----TAAAGAGTCATCAAAAG 121
DB 62 CTGTGTGTAATGTAGCAATAATGGATCAGGTGATTCAAGTAAATAATCGTCATCGAAAG 121

QY 122 ATGGAGTTGAAATCAAGCAGCAAGAGGTACTACGAAAGTACCTTAACACCTTAACGTG 181
DB 122 ATTCTATTAAATTTAAACACGCAATTTAGGCACAACTTAAAGTTCTTAAAGATGCTAAACGTG 181

QY 182 TTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTG 241
DB 182 TCGTAGCGTTGGAATTTTCATTTGTAGATGCTTTAGCGGCACATAAATGTTAAACCTGTTG 241

QY 242 GGATAGCGGATGATAACAAAAAATCGTATTTATTTAAACCAATTAAGAGATAAAATTTGAA 301
DB 242 GGGTTGCTGTAGACAAACAAACCAATCGTATTTATTTAAACCAATTAAGAGATAAAATTTGAG 301

QY 302 AATACACTTCGTAGGAACAGCTAAGCAACCTAACCTTAGAGAAATCAGTAACTTAAAC 361
DB 302 ATTATAAATCTGTTGGTCTCGTAAACCAACCACTTAGAGAAATCAGTAACTTAAAC 361

QY 362 CAGATTTAATTTATTCGTATTAATAGACAAAGGTATTTATTAAGACTTTAAATAAAA 421
DB 362 CAGATTTAATTTATTCGTATTAATAGACAAAGGTATTTATTAAGACTTTAAATAAAA 421

QY 422 TTGCTCTACGATGAACTGAAAGTTTCGATGGAGATTTATAATGAAATATTTGATGCTT 481
DB 422 TTGCTCCAACTTGAATTTAAAGTTTGTGATGGAGATTTATAATGATATAGATGCTT 481

QY 482 TTAACCAATTTCAAAGCTTTAGGTAAAGAGAGAGAGGTTAA 524
DB 482 TTAACCAATTTGCCAAAGCTTTAAATAAAGAGATGATGTAGGTCA 524
```

RESULT 7

```
US-09-815-242-8024
; Sequence 8024, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8024
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(960)
US-09-815-242-8024
```

```
Query Match      26.3%; Score 265.6; DB 9; Length 960;
Best Local Similarity 57.8%; Pred. No. 8.2e-36;
Matches 494; Conservative 0; Mismatches 354; Indels 6; Gaps 1;

QY 132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCCTAAACCGTGTGTGTCT 191
DB 105 AATAAAGATGAATTTAGGAACCTGAAAAAATTAAGAAAAATCCTAAACGTTGTTGTATT 164

QY 192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGATACCGA 251
DB 165 AGAATATATGTTTCTGCTGATTTTTCAGCAGCATTTAGATATGAACCTGTTGTTGTCAGA 224

QY 252 TGATAACAAAAAATCGTATTTTAAACCAATTAAGAGATAAAATTTGAAAAATACACTTC 311
DB 225 TGATGGCAGCAGTAAAAATATAACAACTCAGTAAAGATAGGTTGGGCGCATATGAATC 284

QY 312 TGTAGAACACGTAAAGCAACCTAACTTTAGAGAATAATCAGTAAACCTAAACCCAGATTAA 371
DB 285 GGTGATCTAGACCGCAACCGAATATGGAAGTGTATAGTAAATTTAAACCGGATTTGAT 344

QY 372 TATTGCTGATTAATATAGACACAAAGGTATTTTAAAGACTTTAAAGACTTTAAATTAATTCCTCTAC 431
DB 345 CATAGCAGATGTTAGCAGACATAAGAAAAATCAAAATCAGAAATTAAGCAAAATTTGCTCCGAC 404

QY 432 GATTGAACCTGAAAAAGTTTCGATGGAGATTTATAATGAAAAATTTGATGCTTTTAAAAACAAT 491
DB 405 AATCATGTTAGTTAGCGGTACGGGAGATTTAATGCAAAATTTGATGCAATTTTAAACAGT 464

QY 492 TTCAAAAGCTTTAGGTAAAGAGAGAGGTAAAGAAACGCTTTAGAGAACACGATPAAGAA 551
DB 465 CGCTAAACGCTAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 524

QY 552 AATTGAAGAAATATAAAAAAGAAAAATAACTATATGATTAATAATCAAAAGGTATTTGCTGCGAGT 611
DB 525 ATTACGGGAGATTAGAAAAAGAAAAATTTGAAACAGATACGTTTAAAAAACTGCAATTTGCAATCGG 584

QY 612 AGCTGCTAAATCAGGTTTGTGCTCATCCCAAGCAACTCTTATGTTGCTCAATTCCTTAAG 671
DB 585 TATCTCAAGACGAGGTATGTTTATTAATTAATGAAGATACATTTATGGACCAATTTCTTAAT 644

QY 672 TCAACTAGGTTTTTAAAGAGAGCATTAAGTGTATGTTTACTAAAGGTTTAAAGTAAGTATCT 731
DB 645 TAAATGGGTATTCACCTGAACTCACAACAAAGACAAAACTACGCAATGTTGTTGTAACGCAAA 704

QY 732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCTCTGAGCGTAT 791
DB 705 GGGTGGTCTCTTATATTTTATTTTAAATTAATGAAGAACTTCCCAATATCAATCCAAAAAGTTAT 764

QY 792 GTTCATAATGACAAACAAAGCAAGTCTTAAACGAACCTTCCTCTTAAAGAACTTAGAAAAAGA 851
DB 765 GATTTTACCCACTGACGGAAGAAAAACGACAAAAATAG-----AACGAATATTCATTGATCC 818

QY 852 TCCTGTATGGAAGAAATTTAAACGCTGTGAAAAAATCAACGTTGTTGATATTTTATAGACCGTGA 911
DB 819 TGCAGTTTGGAAATCATTAAAAAGCTGTGAAAGATTAACAAAGTTTATGACGTTTGAACCGAAA 878
```

QY 912 CTTATGGGCAAGTACGCTGGTTTAAATTTCTTTCAGAGAAATGCGAAAGAACTTGTGGA 971
Db 879 TAAGTGGTGAANTCAGGGGTATTATCCCAAGTGAAGATGCGCAGAGATTTAGAAAA 938
QY 972 ATTATCTAAGAAAG 985
Db 939 AATTGCAGAAAAAG 952

RESULT 8

US-09-815-242-4194

; Sequence 4194, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-05-23

; PRIOR FILING DATE: 2000-05-23

; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-11-27

; PRIOR FILING DATE: 2000-12-22

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4194

; LENGTH: 957

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-09-815-242-4194

Query Match 26.2%; Score 264.6; DB 9; Length 957;
Best Local Similarity 57.8%; Pred. No. 1.2e-35;
Matches 493; Conservative 0; Mismatches 354; Indels 6; Gaps 1;

QY 132 AATCAGCAGCAGAGAGGTTACTAGCAAGTACCTAAACACCCCTAAACGCTGTTCTGTCT 191
Db 111 AATAAAGATGAATAGAACTGAATAAATAAGAAAAATCCCTAAACGCTGTTCTGTATT 170
QY 192 TGAGTATTCATTTGTTGATCGCTAGTTGCTTTAGATGTTTAAACCTGTTGGGATACGGA 251
Db 171 AGAATATAGTTTGTCTGATTATTATAGCAGCATTAGATATCAACCTGTTGGTATTGAGA 230
QY 252 TGATAACAAAAAATCGTATTATTATTAACCATTAAGAGATAAAATTTGAAAAATACACTTC 311
Db 231 TGATGGCAGCAGTAAATAATAACAAAGTCAGTAAGAGATATAGGTTGGGGCATATGAATC 290
QY 312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAATCTTAAACACGATTAAAT 371
Db 291 GGTGTTGATCTAGACCGCACCGAATATGGAAGTGATGAAGTAATTAACCCGATTGAT 350
QY 372 TATTGCTGATAATAATAGACAAAGGTTATTTATAAGACTTAAATAAAATTCCTCTAC 431
Db 351 CATAGCAGATGTTAGCAGACATAAGAAAAATCAAAATCAGAAATTAAGCAAAATTCCTCGAC 410

QY 432 GATTGAACTGAAAGTTTCGATGGAGATTATAATGAAAAATATTGATGCTTTTAAAAACAAT 491
Db 411 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAAAATATTGATGCAITTTAAAAACAGT 470
QY 492 TTCAAAAGCTTTAGGTAAGAAGAGAGGTTAAAAAACGCTTAGAAGAACACGATAAGAA 551
Db 471 CGCTAAAGCAGTAGGCAAGAGAAAGAGGCGAGAAACGCTCTGGAAAAGCATGATAAAT 530
QY 552 AATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAATCAAAAGGTTATTCCTGCGAGT 611
Db 531 ATTAGCGGAGATTAGAAAAGAAATTTGAACAGATACGTTTAAAAACTGCAITTTGCATTCCG 590
QY 612 AGCTGCTAAATCAGGTTTGTCTCATCCAAAGCAACTCTTATGTTGTTGCTCAATTCCTAAG 671
Db 591 TATCTCAAGAGCAGGTATGTTTATTATAATGAAGATACATTTATGGGACCAATTTCTTAAT 650
QY 672 TCAACTAGGTTTAAAGAGCATTAAGTGATGATGCTTACTAAAGCTTTTAAAGTAAGTATCT 731
Db 651 TAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACCTACGATGTTGGTGAACGCA 710
QY 732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTCAATCTCTGAGCGTAT 791
Db 711 GGGTGTCTTATATTATTATTAATTAATGAAGAACTTGGCAATATCAATCCAAGATTAT 770
QY 792 GTTCATAATGACAAAACAAAGCAAGTTCTAAACGAACTTCTACTAAAGAACTAGAAAAAGA 851
Db 771 GATTTTAGCCACTGACGGAAGAAACCGACAAAATAG-----ACGAAATTCATTGATCC 824
QY 852 TCCTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGTGTTGATATTATTAGACCGTCA 911
Db 825 TGCAGTTTGGAAATCATTTAAAAAGCTGTGAAGATTAACAAAGTTTATGACGTTGACCGAAA 884
QY 912 CTTATGGGCAAGATCACTGCTGGTTTAAATTTCTTCAAGAGAAATGCGCAAGAACTTGTTCGA 971
Db 885 TAAGTGGTTGAATCAAGGGTATTATTCGAAGTGAAGATGAGCAGAGATTAGAAAA 944
QY 972 ATTATCTAAGAAA 984
Db 945 AATTGCAGAAAAA 957

RESULT 9

US-10-470-048B-301

; Sequence 301, Application US/10470048B

; Publication No. US2005003744A1

; GENERAL INFORMATION:

; APPLICANT: MEINKE ET AL.

; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF

; FILE REFERENCE: SONN:035US

; CURRENT APPLICATION NUMBER: US/10/470,048B

; CURRENT FILING DATE: 2003-07-25

; NUMBER OF SEQ ID NOS: 603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 301

; LENGTH: 957

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-10-470-048B-301

Query Match 26.1%; Score 263.2; DB 21; Length 957;
Best Local Similarity 55.9%; Pred. No. 2.1e-35;
Matches 524; Conservative 0; Mismatches 408; Indels 6; Gaps 1;

QY 48 ATTGTTTGTTTTAAATGCAACTGCGCATGTGGAATAATAGTTTCAAGTAACCTCAAGTAA 107
Db 21 ACTAGTTGTGTTTCATGCTAAATTTTACTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80
QY 108 AGAGTCATCAAAAGATGGAGTTGAATCAAGCAGCAAGAGGCTACTACGAAAGTACCTAA 167
Db 81 TGAAGAGAAATCGAAATGACCGCAATAAAGATGAATTAGGAACCTGGAATAAATTAAGAA 140

```
QY 168 ACACCTAAACGCTGTTGTTGTTCTTGAGTATTCATTTGTTGATCGTTAGTTGCTTTAGA 227
Db 141 AAATCCTAAACGCTGTTGTTGTTATTAGAATATAGTTTGTCTGATTTATTAGCAGCAATTAGA 200
QY 228 TGTTAACCTGTTGGATAGCGGATGATAACAAAAAATCGTATTTATTAACCAATTAAAG 287
Db 201 TATGNAACCTGTTGGTATTGCAGATGATGCGGCACTAAAATATAACAAAGTCAGTAAAG 260
QY 288 AGATAAAATTCGAAATACATCTCTGTAGAACACGTAAGCAACTAACTCTAGAGAAAT 347
Db 261 AGATAAGATTGGGCAATATGAATCGTTGGATCTAGACCGCAACCGAATATGGAAGTGTAT 320
QY 348 CAGTAAACCTTAACCAAGATTTAAATTTCTGCTGATTAATATAGACACAAAGGTATTTATAA 407
Db 321 AAGTAAATTAACACCGGATTTGATCAATGCGAGATGTTAGCAGACATAAGAAAATCAAAATC 380
QY 408 AGACTTAAATTAATTTGCTCTACCAATGAACTGAAATGTTTGAATGAGATTTATAATGA 467
Db 381 AGAATTGAGCAAAATTTGCTCGACAACTCATGTTAGTTCAGCGTACGGGAGATTATAATGC 440
QY 468 AAATATTGATGCTTTTAAACAAATTTCAAAGCTTTAGGTAAAGAAAGAAAGGTAAGAA 527
Db 441 AAATATTGAAGCAATTTAAACAACTGCTTAAAGCAGTAGGCAAGAGAAAGAAAGCGAGAA 500
QY 528 ACGTTAGAAACACGATAGAAATTTGAAGATATATAAAGAAATTAATCTATGATAA 587
Db 501 CGCTCTGAAAGCATGATTAATATTAGCGGAGATTAGAAGAAATTTGAACAGAGTAC 560
QY 588 AAATCAAAAGGATTTGCTCGAGTAGCTGCTAAATCAGGTTTGTCTGCTCAATCCAAAGCAA 647
Db 561 GTTAAATCTGCAATTTGCAATTCGGTATCTCAAGACAGGTATGTTTATTAATATGAAGA 620
QY 648 CTCTTATGTTGGTCAATTTCTTAAGTCAACTAGGTTTTTAAAGAACATTTAAGTGTATGT 707
Db 621 TACATTTATGGACAAATTTCTTAATTTAAATGGGTATTCACCTGAAGTACACAAAAGACAA 680
QY 708 TACTAAAGGTTTAAGTAGATATCTTAAAGGACCTTACTTACAAATGACACCTGAACTTT 767
Db 681 AACTACGCAATGTTGGTGAACGCAAGGGTGGTCTTATATATATTAAATATGAAGAACT 740
QY 768 ATCTCAAGTGAATCTCGAGCGTATGTTTCATATATGACAAACAAAGCAAGTTCCTAACGAACC 827
Db 741 TGCCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGCGGAAAAACGGACAAAA --- 797
QY 828 TTCACATAAAGAACTAGAAAAAGATCTCTGTATGGAAGAAATTTAAACGCTGTGAAAAATCA 887
Db 798 ---TAGAACGAAATTCATTTGATCTCCCTGCAGTTTGGAAATCAITTTAAAGCTGTGAAAGATAA 854
QY 888 ACGTGTTCATATTTTAGACCGTGACTTATGGCAAGATCACGTGGTTTTAATTTCTTCAGA 947
Db 855 CAAAGTTTATGACGTTGACCGGAAATTAAGTGGTTGAATCAAGGGGGAATTCGCAAGTGA 914
QY 948 AGAAATGCAAAAGAACTTTGTTGAATTTATCTAAGAAAG 985
Db 915 AAGTATGCGCAGAGATTTAGAAAAAATTCGCAAAAAAG 952
```

RESULT 10

```
US-10-282-122A-7949
; Sequence 7949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
```

```
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7949
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; US-10-282-122A-7949
```

```
Query Match 26.1%; Score 263.2; DB 17; Length 960;
Best Local Similarity 55.9%; Pred. No. 2.1e-35;
Matches 524; Conservative 0; Mismatches 408; Indels 6; Gaps 1;

QY 48 ATTGTTGTTTAAATTGCAACTGCGACGATGTGGAAATTAATAGTTCAAGTAACCTCAAGTAA 107
Db 21 ATTAGTTGTGTTTATGCTAAATCTTAGTTGTAGCAGTAGCGGGTTGTGCTCAAAAAGATAC 80
QY 108 AGAGTCATCAAAAGATGAGTTGAAATCAAGCACGAAAGGTACTACGAAAGTACCTAA 167
Db 81 TGAAGAGAAACTGAAATGACGACAAATAAAGATGAATTAGAACTGAAAAAATTAAGAA 140
QY 168 ACACCTAAACGCTGTTGTTGTTCTTGAGTATTCATTTGTTGATCGGTTAGTTGCTTTAGA 227
Db 141 AAATCCTAAACGCTGTTGTTGTTATTAGAATAAGTTTGTCTGATTATTTAGCAGCATTAGA 200
QY 228 TGTTAACCTGTTGGGATAGCGGATGATAACAAAAAATCGTATTTATTAACCAATTAAAG 287
Db 201 TATGAAACCTGTTGGTATTGCGAGATGATGCGTGGATCTAGACCGCAACCGAATATGGAAGT 260
QY 288 AGATAAAATTCGAAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
Db 261 AGATAAGATTGGGCAATATGAATCGTTGGATCTAGACCGCAACCGAATATGGAAGTGTAT 320
QY 348 CAGTAAACCTTAACCAAGATTTAAATTTATTTGCTGATAATTAATAGACACAAAGGTATTTATAA 407
Db 321 AAGTAAATTAACACCGGATTTGATCAATTTGAGATGTTAGCAGACATAAGAAAATCAAAATC 380
QY 408 AGACTTAAATTAATTTGCTCTACCAATGAACTGAAATGTTTGAATGAGATTTATAATGA 467
Db 381 AGAATTGAGCAAAATTTGCTCGACAACTCATGTTAGTTCAGCGTACGGGAGATTATAATGC 440
QY 468 AAATATTGATGCTTTTAAACAAATTTCAAAGCTTTAGGTAAAGAAAGAAAGGTAAGAA 527
Db 441 AAATATTGAAGCAATTTAAACAACTGCTTAAAGCAGTAGGCAAGAGAAAGAAAGCGGAGAA 500
QY 528 ACGTTAGAAACACGATAGAAATTTGAAGATATATAAAGAAATTAATCTATGATAA 587
Db 501 CGCTCTGAAAGCATGATTAATATTAGCGGAGATTAGAAGAAATTTGAACAGAGTAC 560
```

Qy	588	AAATCAAAAGGTTATGGCTGCAGTAGCTGCTAAATCAGGTTTGCTTCATCCAGCAA	647
Db	561	GTAAAAATCTGCAATTTGCAATTCGGTATCTCAAGACGAGTATGTTTATTAATAATGAAGA	620
Qy	648	CTCTTATGTTGGTCAATTCCTAAAGTCAACTAGGTTTAAAGAGCAATTAAGTGATGATGT	707
Db	621	TACATTTATGGGCAAAATCTTTAAATTTAAATGGGTATTCAACTGAAGTCACAAAAGACAA	680
Qy	708	TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACCTGAACTTT	767
Db	681	AACTACGCATGTTGGTGAACCAAGGGTGGCTTATATATATTTAAATAATGAAGAACT	740
Qy	768	ATCTCAAGTGAATCTGAGCGTATGTTTCATATATGACAAACAAAGCAAGTCTTAACGAACC	827
Db	741	TGCCAATATCAATCAAAAGTTATGATTTTGGCCTGACGGAACCGACAAAAA	797
Qy	828	TTCACTAAAAGAACTAGAAAAAGATCCGTATGGAAGAAATTTAAACGCTGTGAAAAATCA	887
Db	798	---TAGAACGAAATTCATTGATCTGCAAGTTTGGAAATCAITTTAAAGCTGTGAAAGATAA	854
Qy	888	ACGTGTTGATATTTAGACCGTGACTTATGGCGAAGATCACGTGTTTAAATTTCTTCAGA	947
Db	855	CAAAAGTTTATGACGTTGACCGTACCGGAAATTAAGTGGTTGAAATCAAGGGGGATTCGCAAGTGA	914
Qy	948	AGAAATGGCAAAAGAACTTTGTTGAATTTATCTAAGAAAG	985
Db	915	AAGTATGGCAGAAGATTTAGAAAAAAATTTGCAGAAAAAG	952

```

RESULT 11
US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-355
Query Match 22.1%; Score 222.6; DB 8; Length 668;

```

RESULT 12

US-10-329-624-355
Sequence 355, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSES: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224

```
;
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-10-329-624-355

Query Match      22.1%; Score 222.6; DB 18; Length 668;
Best Local Similarity 74.8%; Pred. No. 1.6e-28;
Matches 279; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 627 TTTCTGCTCATCAAGCAACTCTTATGTTGGTCAATCTCTAAGTCAACTAGGTTTAA 686
Db 1 TTTATTAGCACATCCAAACTATTTCATATGTTGGCAATTTTAAACGAACTAGGATTTAA 60

Qy 687 AGAAGCATTAAAGTATGATGTTACTTAAAGGTTTAAAGTATCTTAAAGGACCTTACTT 746
Db 61 AAATGCATTAAAGTGACGATGTAAACAAAGGTTTAAAGTAAATATTGAAAGGACCTTACTT 120

Qy 747 ACAATGAACACTGAAACTTTTATCTCAAGTGAATCCTGAGCGTATGTTTCAATAATGACAAA 806
Db 121 ACAATTAGACACTGACACTTTAGCTGATTTAAATCCAGAGCGTATGATCATTAATGACAGA 180

Qy 807 CAAGCAAGTCTTAACGAACTTCACTAAAGAACTAGAAAAGATCCTGTATGGAAGAA 866
Db 181 TCATGCTAAAAAGATTCTGCTGAATTCAGAAAGTTTACAAGAAGATGCAACATGGAAGAA 240

Qy 867 ATTAAACGCTGAAATCAACGTTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
Db 241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGGCAAGATC 300

Qy 927 ACGTGGTTTAAATTTCTCAGAAAGAAATGGCAAGAACTTGTGAAATTTATCTAAGAAAGA 986
Db 301 TCGTGGCTTAATTTCTTCTGAAAGAAATGCGCTAAGAACTTGTGAAATTTCAAAAAAGA 360

Qy 987 TAGTAAAAAAGAT 999
Db 361 ACAAAAGTAAGT 373

RESULT 13
US-08-781-986A-2556
; Sequence 2556, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446

;
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-2556

Query Match      16.1%; Score 162.2; DB 8; Length 242;
Best Local Similarity 79.3%; Pred. No. 2.3e-18;
Matches 191; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 208 GATCGGTTAGTGTCTTTAGATGTTTAAACCTCTTGGGATAGCGATGATTAACAAAAAAT 267
Db 2 GATGCATTAGCAGCATTAGACGTTTAAACCACTTGGTATTGCTGATGATGGTAAAGAAAAA 61

Qy 268 CGTATTATTAAACATTAAAGATATAAATTGGAATAATACACTTCTGTAGGAACACGTAAG 327
Db 62 CGTATCATTAACCACTTAGAGAAAAAATTGGGGAATTATACCTTCTGTAGGTACACGTAAA 121

Qy 328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAAGATTTAAATTATTCGTGATAATAAT 387
Db 122 CAGCAAACTTTAGAAGAAATTTAGTAAATTAATAACCGGATTTAATTATCGCTGTAGCAGT 181

Qy 388 AGACACAAAGGATTTTATAAAGACTTAAATAAATAATTTGCTCTCTAGATTGAACCTGAAAA 447
Db 182 AGACATANAGGTATTAATAAAGAAATTAANCAAAATTTGACCAACATTTATCATTTAAAGAGT 241

Qy 448 T 448
Db 242 T 242

RESULT 14
US-10-329-624-2556
; Sequence 2556, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
```

```

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30706
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-132A-30706

```

	Query Match	15.5%;	Score 156.4;	DB 17;	Length 897;
	Best Local Similarity	51.9%;	Pred. No. 3.2e-17;		
	Matches 454;	Conservative 0;	Mismatches 381;	Indels 39;	Gaps 3
Qy	123	TGGAGTTGAAATCAAGCACGAAGAAGTACTACGAAAGTAGTACCTAAACACCCCTAAACGTTG	182		
Db	63	TGCTGTTTACCGTCAAGATCAAAAGAGGGAATTTACCTTAGATAGCGTACCTTAAACGTTG	122		
Qy	183	TGTTGTTCTTTGGAGTATTCAATTTGTTGATGGTTAGTTAGTCTTTAGATGTTTAAACCTGTTGG	242		
Db	123	TGTTGCCCTTGAATATTCTTATGTGGATGCACTTGCACAAAATTGTTGTGACGCCCTGTTGG	182		
Qy	243	GATAGCGGATGATAACAAAAAATTCGTATTATTAAACCATTTAAGAGATAAAATTTGGAAA	302		
Db	183	TGTCGCAGATGCACAATGATAAAACCCGTATTCTGCNAAAAGTACGCGATAAAGTGCAGCC	242		
Qy	303	ATACACTTCTGTAGGAACAGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTTAAACC	362		
Db	243	ATGGGAATCAGTGGGAGCGCTTCTCAACGAGTTAGAGCGATTCTGCACTTTAAACC	302		
Qy	363	AGATTTTAATTATTCCTGATATAATATAGACACAAAGGTTATTATAAAGACTTAATAATAAT	422		

423	Qy	TGCTCCTACGATTGAACTGAAAGCTTTTCGATGGAGATTATTAATGAAAAATTTATGATGCTTT	482
363	Db	CGCGCCGACAGTCGTCGTTTAATCCGCGCATGAGAACTATCAAGAAAAACCTTTGAAACCGC	422
483	Qy	TAAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGGTAAAAAACGCTTAGAAGAAACA	542
423	Db	ACAAAAAATCGGTGATTATTATAGGTAAATCAAAAGAAATCAAGAGCGGTATTTCGCAAAACA	482
543	Qy	CGATAAGAAAAATCGAAGAAATATAAADAAGAAATACCTATGGATAAAATCAAAAGCTATT	602
483	Db	TAAGCAGGATATCGGG-----ACATCGCAAAACGTTACGAAAGGAAAAAAGCGAT	536
603	Qy	GCCTGCAGTAGCTGCTAAATCAGGTTTGCTTCTCATCCAAGCAACTCTTATGTTGGTCA	662
537	Db	TATCGGTGTTTCACGTGAACCCCAATTTAAATTTATATAATAGCGAATCTATGCTGGTGG	596
663	Qy	ATTCTTAAGTCAACTAGGCTTTTAAAGAGACATTAGTGATGATGTTTACTTAAAGGTTTAAAG	722
597	Db	CTTAGTGGAGTGTCTAGGTATCAATGCCAAAGCCCGTGCCGATAACCAACCTAATGC	656
723	Qy	TAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCC	782
657	Db	TTCGTGTGGTTTAGAACAG-----TGGCGGCAGAAAAAGCC	692
783	Qy	TGAGCGTATGTTTCATATGACAAACAAAGCAAGTTCTTAACGAACCTTCACTATAAAGAACT	842
693	Db	TGATCTGATGATCTTAATCCATTA-----TCGTGATGAGATATTGCAAGAAATG	743
843	Qy	AGAAAAAGATCCTGTATGGAAAGAAATTAAACGCTGTGAAAAATCAACGCTGTTGATATTTT	902
744	Db	GGAAAAAGAACCTTATGAAAAATTTATCCCTGCGGTAAAAAATGGTCAAGTCATCTTAGC	803

QY 903 AGACCGTACCTATGGCAAGTACAGTGGTTTAATTTCTTCAGAGAAATGGCAAGA 962
DB 804 TAATGATAAATTTATGGCAAGACGCGTGGTATTGATCCGCTGAAGTAAATGGCTAAAGA 863
QY 963 ACTTGTTCGAATTATCTAAGAAAGATAGTAAAAA 996
DB 864 AGTCCAGACTTTGTAAACGAATCCGCCAATAA 897

RESULT 16

US-09-974-300-2066
; Sequence 2066, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2066
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2066

Query Match 14.8%; Score 149.4; DB 9; Length 801;
Best Local Similarity 51.8%; Pred. No. 4.8e-16;
Matches 414; Conservative 0; Mismatches 376; Indels 9; Gaps 3;

QY 176 AACGTGTTGTTCTTCTAGTATTTCATTTGTTGATCGGTAGTTCGTTTAGATGTTAAAC 235
DB 1 AAAAAAGTTTGTGTTCTTGAATTTGGGATTTATTGATCGCTCCTTGATCGGGTATTAAAGC 60
QY 236 CTGTTGGGATACGATGATACAAAAAATCGTATTATTAAACCATTAAGAGATAAAA 295
DB 61 CTGTCGGAATTCGGACGACGGCAACCTAAGTTTATTACGAGAGGTACGGGAAAAA 120
QY 296 TTGGAATAATACATCTCTGTAGGAACACGTAAGCAACCTAATTTAGAGAAATCAGTAAAC 355
DB 121 TCAAGGGGTATCTTCAGTCGGTTCGCGCGCCAGCAAGCTTTGAAAAAATTTGCTTCTT 180
QY 356 TTAACACAGATTAAATTTGCTGATATAATAGACACAAAGGTATTATTAAAGACTTAA 415
DB 181 TAAAGCCGATTTAAATTTGCGGATTCGAGCAGGACACAGCGCTTATGATAGCTGT 240
QY 416 ATAAATGCTCTCAGATTGAACGTGAAGTTTCGATGGAGATATATGAATAATATG 475
DB 241 CGAAATTTGCGCGCAATTCGCGCTCAAAAATTTGAATGCCGATTTATCAGGACGCTTG 300
QY 476 ATGCTTTTAAAAAATTTTCAAAAGCTTTAGTAAAGAAAGAAAGGTAAAAAACGCTTAG 535
DB 301 ATGCATCTCTTACGATTGGAAGCGCTCGGCAAGAAAGCGCAATGGAGAAAAAATGG 360
QY 536 AAGAACAGATGAAGAAATTTGAAGATATATAAAGAAATACTATGGTAAATCAAA 595
DB 361 CTGAACATAAAGAAAGCTGGATGAATTTGAAAACAGAAATTCGCGACGCGGAAACAGACA 420
QY 596 AGGTATTGCTCAGTAGCTGCTAAATCAGGTTTGTCTCATCAAGCAACTCTTATG 655
DB 421 TTTCTTCTCGGACACAAATGAAGAAATCACCGTGGCGATGNA---AACTTTTCA 477
QY 656 TTGTCCTTCTTAAAGTCAACTAGGTTTTPAAGAGCATTAAGTGAATGTTACTAAAG 715
DB 478 CGTCTCAGCTTCTGACGAAAAATCGGCTATACATACGTTGTCGAGACAGCG---GCAAG 534

QY 716 GTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTTATCTCAAG 775
DB 535 GCGATGCCGAAAAACGGTGAATCCGTCATATTAATAATGACCGCTGAAACAGCTGCTTGAGA 594
QY 776 TGAATCTCTGACGCTATGTTTCAATAGC---AAACAAAGCAAGTTCTTAACGAAACCTTCAAC 832
DB 595 AAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAGTCGATGAAGACGGGA 654
QY 833 TAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAACGCTGTGAAAAATCAACGTG 892
DB 655 AAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCAAGG 714
QY 893 TTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAAGAAA 952
DB 715 TTTATGAAGCCGACAGATTCCGCTCGCTCCGACGAGCATTCACGAGCAGATGAGC 774
QY 953 TGGCAAAAAGAACTTTGTTGA 971
DB 775 TGATGATGAATCGATCA 793

RESULT 17

US-10-282-122A-9415
; Sequence 9415, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9415
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Bacillus anthracis
US-10-282-122A-9415

Query Match 14.2%; Score 143.2; DB 17; Length 972;
Best Local Similarity 55.0%; Pred. No. 5.7e-15;

	Matches	308;	Conservative	0;	Mismatches	243;	Indels	9;	Gaps	1
Qy	25	AAAAATTTAAAGTAAATGGCTTAATGTTGCTGTTTAAATTCGAACACTGCAGCATGTGGAAT	84							
Db	7	AAAAATTCCTCAGTATTTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA	66							
Qy	85	AATAGTTCAGTAACACTCAAGTAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCACGAA	144							
Db	67	GAGGAGAAAAAAGAAACAAAAGCGGACAAATAAAAATCAAGCTATAACAATTAATAACACGCT	126							
Qy	145	GAAGGTACTACGAAAGTACCTAAACACCCCTAAACCGTGTGTGTTGTTCTTGAGTATTCATTT	204							
Db	127	GAAGGGGAAACGAGTTAGATTAACACGCGGAAAAAAGTAGTTGTACTTGAATGGGTATAT	186							
Qy	205	GTTGATGCGTTAGTTGCTTTTAGATGTTTAAACCTGTGGGATAGCGGATGATAACAAAAAA	264							
Db	187	TCAGAAGACTTATTAGCACTTTGGTGTTCAGCCAGTAGGATGCGCAGACATTAAAGAAATAT	246							
Qy	265	AATCGTATTATTAACCACTTAAGAGTAAATTTGGAAAATACACTTCCTGTAGGACACAGT	324							
Db	247	AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACAGT	306							
Qy	325	AAGCAACCTCACTTAGAAGAAATCAGTAAACCTTAAACCAGATTTAAATTTATTCCTGATAT	384							
Db	307	CAACAAACCAACTTAGAAGAAATTAGCCGTTTAAACCAGATTTAAATTTATCACAGCTCA	366							
Qy	385	AATAGACACAAAGGTATTTTAAAGACTTAAATAAAATGTCCTCACGAT	435							
Db	367	TTCCGTGGTAAAGCAATTTAAAAATGAAATAGAACAAATTTGCAACCAACAGTTATGTTTGTAT	426							
Qy	436	GAACTGAAAGTTTCGATGGAGATTTAATGAAATATTGATGCTTTTAAACAAATTTCA	495							
Db	427	CCATCAACACGCAATTAACGATCCTTTGCTGAAATGACAGAAACATTTTAAACAAATTTGCA	486							
Qy	496	AAAGCTTTAGGTAAAGAAAGAAAGGTAATAAAAAACGCTTAGAAGAACACGATTAAGAAAAAT	555							
Db	487	AAAGCAGTTGGAAAGAAAGAAAGAGGTAAAAAAGTATTAGCTGATATGATGAAGCAATTC	546							
Qy	556	GAAGAATATAAAAAAGAAAT	575							
Db	547	GCTGATGCAAAAGCAAAAAT	566							

RESULT 18'

```

US-09-815-242-2107/c
; Sequence 2107, Application US/09815242
; Patent No. US20020061569A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Onisen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
;
; TITLE OF INVENTION: Identification of
; TITLE OF INVENTION: Prokaryotes
;
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815
; CURRENT FILING DATE: 2001-03-21
;
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
;
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
;
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
;
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
;
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
;
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
;

```

```

; PRIOR APPLICATION NUMBER: 60/269, 308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2107
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-2107

Query Match          12.7%; Score 128; DB 9; Length 321;
Best Local Similarity 67.8%; Pred. No. 1.7e-12;
Matches 179; Conservative 0; Mismatches 85; Indels 0; Gaps 0

Qy      132  AATCAAGCAGCAAGAGGTACTGAAAGTACCTAAACACCCCTAAACCGTGTGTGTGTCT 191
      |||
Db      271  AATAAAGATGAATTAGAACTGAAAAAATTAGAAAAATCCTAAACGTGTGTGTATT 212

Qy      192  TGAGTATTCAATTTGCTGATCGGTTAGTGTGCTTTAGATGTTAAACCTGTGGGATAGCGGA 251
      |||
Db      211  AGAATAAGTATTTGCTGATTATTTAGCAGCATTAGATATGAACCTGTGTGGTATTGCAGA 152

Qy      252  TGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGCAAAATACACATTC 311
      |||
Db      151  TGATGGCAGCACTAAAAAATATACAAGTCTAGTAAGAGATAAGATTGGGGCATATGAATC 92

Qy      312  TGTAGGAAACAGTGAAGCAACCTAACTTAGAGAAATCAGTAAACCTTAAACACGAGATTTAAT 371
      |||
Db      91  GTTTGGATCTAGACCGCAACCGATATGGAAGTGATAGTAAATTTAAACACCGATTGAT 32

Qy      372  TATTGCTGATAAATAATAGACAA 395
      |||
Db      31  CATTCAGATGTTAGCAGACATAA 8

```

RESULT 19

US-10-282-122A-4605/C
; Sequence 4605, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636

```
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4605
/ LENGTH: 321
/ TYPE: DNA
/ ORGANISM: Staphylococcus aureus
US-10-282-122A-4605

Query Match      12.7%; Score 128; DB 17; Length 321;
Best Local Similarity 67.8%; Pred. No. 1.7e-12;
Matches 179; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 132 AATCAAGCAGCAAGAGGCTACTACGAAAGTACCTAAACACCCCTAAACCGTGTGTGTCT 191
DB 271 AATAAAGATGATTTAGGAATCTGAATAAATTAAGAAATCTTAACCGTGTGTGTATT 212
QY 192 TGAGTATTCTTTTGTGATGGCTTAGTTGCTTTAGATGTTAAACCTGTGGATAGCGGA 251
DB 211 AGAATATAGTTTGTGCTGATTTATTTAGCAGCATTAGATATGAACCTGTGTGTGTCAGA 152
QY 252 TGATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATGGAAATACACTTC 311
DB 151 TGATGGCAGCACTAAAAATATAAACAAGTCTAGTAAGAGATAGATTTGGGGCATATGAATC 92
QY 312 TGATGAACACCTAGCAACCTTAACCTTAGAAGAAATCAGTAAACCTTAACACAGATTTAAT 371
DB 91 GGTGGATCTAGACCGCAACCGAATATGGAAGTATAGTAATTAATTAACCGGATTTGAT 32
QY 372 TATTGCTGATATAATAGACACAA 395
DB 31 CATTGCAGATGTTAGCAGACATAA 8
```

RESULT 20

```
US-08-781-986A-604
/ Sequence 604, Application US/08781986A
/ Publication No. US20030054436A1
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5255
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,986A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 604:
/ SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 2115 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-781-986A-604
```

Query Match 12.4%; Score 125.4; DB 8; Length 2115;

Best Local Similarity 53.2%; Pred. No. 7.5e-12;

Matches 289; Conservative 1; Mismatches 247; Indels 6; Gaps 1;

```
QY 458 ATTATAATGAAAAATATTGATGCTTTTAAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
DB 1 ATTATAATGCAATATTGAAGCAATTTAAAAACAGTCGCTAAAGCAGTAGTGCAAGAGAAAAG 60
QY 518 AAGGTAAAAACGCTTAGAAGAACACGATAGAAAAATTTGAAGAATATATAAAAAGAAATAA 577
DB 61 AAGCGAGAGACGCTGTGAAAAAGCATGATAAAATATTAGCGGAGATTAGAAGAAATTTG 120
QY 578 CTATGGATAAAAATCAAAAAGCTATTGCTCGCAGTAGCTGCTAAATCAGGTTTGTCTGCTC 637
DB 121 AACAGATGACGTTTAAATCTGCATTTTGCATTCGGTATCTCAAGACGAGGTATGTTTATTA 180
QY 638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTTAAGTCAACTAGGTTTAAAAAGAGCATTTAA 697
DB 181 ATAATGAAGATACATTTTATGGACAAATTTCTTAATTAATGATGGTATTCAACCTGAAGTCA 240
QY 698 GTGATGATGTTTACAAAGGTTTAAAGTAAAGTATCTTTAAAGGACCTTTACTTACAATGAACA 757
DB 241 MAAARAAMAAAACCTACGCATGTTGGTGAACCAAGGGTGGTCTTATATATATATTTAAATA 300
QY 758 CTGAACCTTTATCTCAAGTGAATCTGAGCGTATGTTTCATATGACAAACAAACCAAGTT 817
DB 301 ATGAAGAACTTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCCACTGACGGAANAACGG 360
QY 818 CTAACGAACCTTCACTAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTTAAACGCTG 877
DB 361 ACAAAA-----TAGAACGAAATTCATTGATCTCGCAGTTTGGAAATCATTTAAAGCTG 414
QY 878 TGAATAATCAACGTTGTGATATTTTAGACCGTGACTTATGGGCAAGATCACTGTTTAA 937
DB 415 TGAAGAATAACAAAGTTTATGACGTTGACCGAAATTAAGTGGTGAATCAAGGGGGGATTA 474
QY 938 TTTCTTCAGAGAATGCGAAAAAGAACCTTTGTTGAATATCTTAAGAAAGATAGTAAAAAAG 997
DB 475 TCGCAAGTGAAGATATGCGAGAAGATTAGAAAAAATTCGAGAAAAACCAAAATAAAAAT 534
QY 998 ATA 1000
DB 535 ACA 537
```

RESULT 21

```
US-10-329-624-604
/ Sequence 604, Application US/10329624
/ Publication No. US20040043037A1
/ GENERAL INFORMATION:
```

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
FILING DATE: 27-Dec-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-10-329-624-604

Query Match 12.4%; Score 125.4; DB 18; Length 2115;
Best Local Similarity 53.2%; Pred. No. 7.5e-12;
Matches 289; Conservative 1; Mismatches 247; Indels 6; Gaps 1;
QY 458 ATTATAATGAAATATGATGCTTTTAAACAAATTTCAAAGCTTTTAAAGAGAAAG 517
DB 1 ATTATAATGAAATATGATGCTTTTAAACAAATTTCAAAGCTTTTAAAGAGAAAG 60
QY 518 AAGGTAAAAAGCTTTAGAGAACACGATAGAAATTTGAAGATATATAAAGAAATAA 577
DB 61 AAGGCGAGAGGCTTGGAAAGCATGATTAATATTAGCGGAGATTAGAAGAAATTTG 120
QY 578 CTATGATAAAAATCAAAAGGATTTGCTGCGATAGCTCTAAATCAGGTTTCTTGCTC 637
DB 121 AACAGAGTACGTTAAATCTGCATTTGCAATTCGGTATCTCAAGAGCGGTATGTTATTA 180
QY 638 ATCCAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTAAAGAGCAATTA 697
DB 181 ATAATGAAGATACATTTATGGGACCAATTTCTAATTAATGGTATTTCAACCTGAATCA 240
QY 698 GTGATGATGTTACTAAGGTTTAAAGTAAATCTTTAAAGCACTTACTTACAAATGAACA 757
DB 241 MAARAARMAAACTACGCAATTTGGTGAACGCAAGGTTGCTCTATATATATTTAAATA 300
QY 758 CTGAACCTTTTACTCAAGTGAATCTCGAGCGTATGTTCAATATGACAAACAAAGCAAGTT 817
DB 301 ATGAAGAACTTGCCAAATATCAATCCAAAAGTTATGATTTTGGCCACTGACGGAACCG 360
QY 818 CTAACGAACCTTCACTTAAAGAACTAGAAAAGATCTCTGTATGGAAGAAATTTAAACGCTG 877
DB 361 ACAAAA-----TAGACGAATTCATTGATCTCGAGTTTGGAAATCATTAAAGCTG 414
QY 878 TGAATAATCAACGTTGTGATATTTTAGACCGTGACTTTAGGGCAAGATCAAGCTGGTTTAA 937
DB 415 TGAAGATTAACAAAGTTTATGAGTTTGACCGAAATTAAGTGGTTGAAATCAAGGGGATTA 474
QY 938 TTCTTTCAGAGAAATGGCAAGAACTTTGTAATTTATCTAAGAAAGATAGTAAAGAG 997
DB 475 TCGCAAGTGAAGATATGGCAGAGATTTAGAAAAAATTCAGAAAAAGCAAAATATAAAT 534
QY 998 ATA 1000
DB 535 ACA 537

RESULT 22
US-10-470-048B-318
; Sequence 318, Application US/10470048B
; Publication No. US2005003744A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: ANTIGENS TO A SPECIFIC PATHOGEN
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 318
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-470-048B-318
Query Match 12.3%; Score 124; DB 21; Length 990;
Best Local Similarity 50.7%; Pred. No. 1.1e-11;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
QY 27 AATTTAAAGTGAATTTGGCTTATCTGTTTAAATTTGCAACTGCAGCATGTGGAATAA 86
DB 12 AATTAATGCTTGTGTTTACGCTTCTCTACTTGTCTTGTAGCAGGATGTAGTGGAA 71
QY 87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA 146
DB 72 TTCAATAAACAATCATCTGATAACAAAGATAAGGAACAACCTCAATTTAAACATGCAAT 131
QY 147 AGGTACTAGGAAGTACCTAAACACCTTAAAGCTGTTGTTCTTGTAGTATTTCTTTGT 206
DB 132 GGGTCAACTGAAATTTAAAGGGAAACCAAGCGTGTGTTACGCTATATCAAGGTGCCAC 191
QY 207 TGATCGGTTAGTGTCTTTAGATGTTAAACCTGTCGGATAGCGGATGATAACAAAAAAA 266
DB 192 TGACCTCGCTGATCTTTAGGTGTTAAACCTGAGTGTCTGAGATCATGACACAAAA 251
QY 267 TCGTATTATTAAACCATTAAGAGATAAAATTTGAAAAATACACTTCTGTAGGAACAGTAA 326
DB 252 ACCGAAATTCGAATACATAAATAAATGATTTAAAGATACATAAGATTGTAGGTCAAGAAC 311
QY 327 GCAACTAATCTAGGAAGAAATCAGTAAACTTAAACGAGTTTAAATTTCTGTGATATAA 386
DB 312 TGCACCTAATCTAGAGGAAATCTCTAAATTTAAACCGGACTTAAATTTGTCGCTCAAAAGT 371
QY 387 TAGACACAAAGGTATTTATTAAGACTTAAATAAAATTTGCTCTACGATTGNACTGAAAA 446
DB 372 TAGAAATGAAAAGTTTACGATCAATTTCTAAATTCGACCC-----AACAG 418
QY 447 TTTTCATGGAGATTATTAATGAAAAAT-ATTGATGCTTTTAAAAACAATTTCAAAGCTTTAG 505
DB 419 TTTCTACTGATACAGTTTCAATTTCAAGAATACAACTAAAGTTAATGGGAAAGCTTTAG 478
QY 506 GTAAAGAGAAAGAGGTAAAAACCGCTTAGAAGAACACGATTAAGAAAAATTTGAAGATATA 565
DB 479 GGAAGAAAAAGAGCTGAAGATTTACTTTAAAGTACGATGATAAAGTACGTGCTCC 538
QY 566 AAAAGA---AATACTATGGTAAAAATCAAAAGGATTTGCTTCAGTGTGCTTAAAT 622
DB 539 AAAAGATGCAAAAGCAAGTATAAGATGCAATGTCATTTGAAAGCTTCAGTTGTTAACT 598
QY 623 CAGGTTTGTCTCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTCACTCAACTAGGTT 682
DB 599 TCCGT---GCTGATCATACAGAAATTTATGCTGGTGGATATGCTGGTGAATCTTAAATG 655
QY 683 TTAAGAAGCAATTAAGTATGATGTTACTAAAGGTTTAAAGTATCTTTAAAGGACCTT 742
DB 656 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAAGTTGATTAATGTTAAAGATATA 715
QY 743 ACTTACAATGAACACTGAACTTTTCTCAAGTGAATCTCTGAGCGTATGTTCAATATGA 802

Db 716 TCCAACTTACATCAAGAAAGCATTCATTAATGAACGCTGATCATATTTTGTAGTAA 775
Qy 803 CAAACAAAGCAAGTTCTTAACGAACCTTCACTA 834
Db 776 AATCAGATCCAAATGCGAAGATGCTGCATTA 807

RESULT 23
US-10-278-946-15
; Sequence 15, Application US/10278946
; Publication No. US20030153733A1
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-278-946-15

Query Match 12.3%; Score 124; DB 16; Length 999;
Best Local Similarity 50.7%; Pred. No. 1.1e-11;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

Qy 27 AATTTAAAGTGAATGGCTTATTGTTTAAATGCAACTGAGCATGCGAAATAA 86
Db 18 AATTAATAATGCTTGTGTAGCTTTCCTACTTGTGTAGCAGGATGAGTGGAA 77

Qy 87 TAGTTCAAGTAACCTCAAGTAAGATGATCAAAAGATGGATGTAATCAAGCAAGA 146
Db 78 TTCAAATAACAATCATCTGATAAAGATAAGGAAACAACCTTCAATTAACATGCAAT 137

Qy 147 AGGTACTACGAAGTACCTAAACACCCCTAAACGCTGTTGTTCTTGTAGTATTCATTGT 206
Db 138 GGGTCAACTGAAATTAAGGGAACCAAGCGTGTGTAGCGCTATATCAAGGTGCCAC 197

Qy 207 TGATCGGTAGTGTGCTTTAGATGTTAAACCTGTTGGATAGCGGATGATAACAAAAA 266
Db 198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGTGTGCTGAGTAATCATGACACAAA 257

Qy 267 TCGTATATTAACCAATTAAGAGATAAATGCGAAATAACATCTCTGTAGGAACAGTAA 326
Db 258 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATACATAAGATGTTAGGTCAAGAAC 317

Qy 327 GCAACCTACTTAGAAGAAATCAGTAAACCTTAAACAGATTTAAATTTATGCTGATAATA 386
Db 318 TCACCTTAATCTTAGAGGAATCTTAAATTAACCGGACTTAATTTGTCGCTCAAGT 377

Qy 387 TAGACACAAAGGTATTTATAAGACTTAAATAAAAAATGCTCTCGATTTGAACCTGAAAAG 446
Db 378 TAGAAATGAAAAAGTTTACGATCAATTAATCTAAATGCGACC-----AACAG 424

Qy 447 TTTTCATGAGATTAATAAGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
Db 425 TTTTACTGTACAGTTTTTCAAAATTCAAAGATACAACTAAGTTAATGGGGAAGCTTTAG 484

Qy 506 GTAAAGAAAGAAAGTAAAAAAGCGTTAGAGAAACACGATTAAGAAAAATTTGAAGAATATA 565

Db 485 GGAAAGAAAAAGAGCTGAAGATTTACTTTAAAAAGTACGATGATAAAAGTAGCTGCATTCC 544
Qy 566 AAAAAGA--AAATAACTATGGATAAAAAATCAAAAGGATTTCCTCGCAGTAGCTGCTAAAT 622
Db 545 AAAAAGATGCAAAAGCAAGATATAAAGATGATGCGCAATTTGAAAGCTTCAGTTGTTAACT 604

Qy 623 CAGGTTTGTCTTCATCCCAAGCAACTCTTATGTTGGTCAATTCCTCAATCAAGTCAACTAGGTT 682
Db 605 TCCGT--GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy 683 TTAAGAAGCAATTAAGTATGATGTTACTAAAGGTTTAAAGTAAAGTATCTTAAAGACCTT 742
Db 662 ATTTAGGATTCAAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGCTAAAGATATTA 721

Qy 743 ACTTACAATAACCACTGAACCTTTATCTCAAGTGAATCCTCGAGGATGTTTCATAATGA 802
Db 722 TCCAACTTACATCTAAAGAAAGCAATTCATTAATGAACGCTGATCATATTTTGTAGTAA 781

Qy 803 CAAACAAAGCAAGTTCTAAACGAACCTTCACTA 834
Db 782 AATCAGATCCAAATGCGAAGATGCTGCATTA 813

RESULT 24
US-10-967-189-15
; Sequence 15, Application US/10967189
; Publication No. US20050053995A1
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD2
; CURRENT APPLICATION NUMBER: US/10/967,189
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: US 10/278,946
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-967-189-15

Query Match 12.3%; Score 124; DB 21; Length 999;
Best Local Similarity 50.7%; Pred. No. 1.1e-11;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;

Qy 27 AATTTAAAGTGAATGGCTTATTGTTTAAATGCAACTGAGCATGCGAAATAA 86
Db 18 AATTAATAATGCTTGTGTAGCTTTCCTACTTGTGTAGCAGGATGAGTGGAA 77

Qy 87 TAGTTCAAGTAACCTCAAGTAAGATGATCAAAAGATGGATGTAATCAAGCAAGA 146
Db 78 TTCAAATAACAATCATCTGATAAAGATAAGGAAACAACCTTCAATTAACATGCAAT 137

Qy 147 AGGTACTACGAAGTACCTAAACACCCCTAAACGCTGTTGTTCTTGTAGTATTCATTGT 206
Db 138 GGGTCAACTGAAATTAAGGGAACCAAGCGTGTGTAGCGCTATATCAAGGTGCCAC 197

Qy 207 TGATCGGTAGTGTGCTTTAGATGTTAAACCTGTTGGATAGCGGATGATAACAAAAA 266
Db 198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGTGTGCTGAGTAATCATGACACAAA 257

Qy 267 TCGTATATTAACCAATTAAGAGATAAATGCGAAATAACATCTCTGTAGGAACAGTAA 326
Db 258 ACCGAAATTCGAATACATAAAAAATGATTTAAAGATACATAAGATGTTAGGTCAAGAAC 317

Qy 327 GCAACCTACTTAGAAGAAATCAGTAAACCTTAAACAGATTTAAATTTATGCTGATAATA 386
Db 318 TCACCTTAATCTTAGAGGAATCTTAAATTAACCGGACTTAATTTGTCGCTCAAGT 377

Qy 387 TAGACACAAAGGTATTTATAAGACTTAAATAAAAAATGCTCTCGATTTGAACCTGAAAAG 446
Db 378 TAGAAATGAAAAAGTTTACGATCAATTAATCTAAATGCGACC-----AACAG 424

Qy 447 TTTTCATGAGATTAATAAGAAAT-ATTGATGCTTTTAAACAAATTTCAAAAGCTTTAG 505
Db 425 TTTTACTGTACAGTTTTTCAAAATTCAAAGATACAACTAAGTTAATGGGGAAGCTTTAG 484

Qy 506 GTAAAGAAAGAAAGTAAAAAAGCGTTAGAGAAACACGATTAAGAAAAATTTGAAGAATATA 565

TELECOMMUNICATION INFORMATION:									
; TELEPHONE: (301) 309-8504									
; TELEFAX: (301) 309-8512									
; INFORMATION FOR SEQ ID NO: 238:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 3775 base pairs									
; TYPE: nucleic acid									
; STRANDEDNESS: double									
; TOPOLOGY: linear									
US-08-781-986A-238									
Query Match 12.3%; Score 124; DB 8; Length 3775;									
Best Local Similarity 50.7%; Pred. No. 1.5e-11;									
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;									
QY	27	AATTTTAAGTGTAAATTTGGCTTATTTGTTTGTGTTTAAATTCGCAACTGCAGCATGTGGAATAA	86						
DB	60	AATTAATAATGCTGTTGTTAGCTTCTCTACTTGTCTTTAGCAGGATGTAGTGGNA	119						
QY	87	TAGTTCAAGTAACTCAAGTAAAGAGTCAATCAAAAGATGGAGTTGAAATCAAGCAGCAAGA	146						
DB	120	TTCAATAAACAATCATCTGATAACAAAGATAAGGAAACAACCTTCAATTAAACATGCAAT	179						
QY	147	AGGTAAGTAAAGTACCTAAACACCCCTAAACGCTGTTGTTCTTGTAGTATTCATTTGT	206						
DB	180	GGGTACAACTGAAATTTAAAGGGAAACCAAGCGCTGTTGTACGCTATATCAAGGTGCCAC	239						
QY	207	TGATCGCTTGTGTTTGTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAA	266						
DB	240	TGACGTCCTGTATCTTTAGGTTTAAACCTGTAGTGTCTGTAGATCATGGACACAAA	299						
QY	267	TCGTATTATTAAACCATTAAGAGATAAAATTTGAAAAATACACTCTGTAGGAAACAGTAA	326						
DB	300	ACCGAAATTCGAATACATAAAAAATGATTTAAAGATATAAGATTGTAGGTCAAGAAC	359						
QY	327	GCAACTTAATTAAGAAATTCAGTTAAACTTAAACAGATTTAAATTTTGTGATATAA	386						
DB	360	TGCACCTTAATTAAGAAATTCCTTAAATTTAAACCGGACTTTAAATTTGTGCGCTCAA	419						
QY	387	TAGACACAAAGGTATTTTATAAGACTTAAATAAAATTTGCTCTACGATTGAACTGAAA	446						
DB	420	TAGAAATGAAAAGTTTACGATCAATTTCTAAATTCGACC-----AACAG	466						
QY	447	TTTCGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAAAACAATTTCAAAGCTTTAG	505						
DB	467	TTTCTCTGTATACAGTTTCAAATTTCAAAGATACAACTAAAGTTAATGGGAAAGCTTTAG	526						
QY	506	GTAAAGAAAGAAAGTAAAAACCGCTTAGAAGAACACGATAGAAATTTGAAGATATA	565						
DB	527	GGAAAGAAAAAGAGCTGAAGATTTTACTTTAAAGTACGATGATAAGATGCTGCATTC	586						
QY	566	AAAAAGA---AATACTATGGATAAATCAAAGGTATTCCTGCAGTAGCTGCTAAAT	622						
DB	587	AAAAAGATGCAAAAGCAAGTATAAGATGCAATGGCCATTGAAAGCTTCAGTTGTTAACT	646						
QY	623	CAGGTTTGTCTCTCATCCAAACCAACTCTTATGTTGGTCAATTCCTTAAGTCAACTAGGTT	682						
DB	647	TCCGT---GCTGATCATACAAAGATTTATGCTGGTGGATATGCTGTGAAATCTTAAATG	703						
QY	683	TTAAAGAAAGCATTAAGTATGATGTTACTAAAGGTTTAAAGTATCTTAAAGGACCTT	742						
DB	704	ATTAGGATTCAAACGCTAATAAAGACTTACAAAAACAAGTTGATATATGTTAAAGATATA	763						
QY	743	ACTTACAAATGAACACTGAACTTTTATCTCAAGTGAATCCTCAGCGCTATGTTTCATATGA	802						
DB	764	TCCAACTTACATCTAAAGAAAGCATTCCTAATTAATGAACGCTGATCATATTTTGTAGTAA	823						
QY	803	CAAAACAAAGCAAGTTCTTAACGAACTTTCACTA	834						
DB	824	AATCAGATCCAAATGCGAAGATGCTGCATTA	855						
RESULT 26									

QY	267	TCGTATTATTAAACCATTAAGAGATAAAATTCGAAATAACACTTCTGTAGGAACACGTAA	326
DB	258	ACCGAAATTCGAATACATATAAAATGATTTAAAGATACTAGATGTAGGTCAAGAACC	317
QY	327	GCAACCTAACTTAGAAGAAATCAGTAAACCTTAAACCGAGATTTAAATTTGCTGATAATA	386
DB	318	TGCACCTTAATTTAGAGAAATCTCTAAATTTAAACCGGACTTAAATTTGTGCGGTCAAAGT	377
QY	387	TAGACACAAAGGTATTATTAAAGACTTAAATTAATTAATTTGCTCTACGATTTGAAGTAAAG	446
DB	378	TAGAATGAAAAGTTTACGATCAATTTATCTAAATTCGACC-----AACAG	424
QY	447	TTTCGATGGAGATTATAATGAAAT-ATTGATGCTTTTAAAAACAATTTCAAAGCTTTAG	505
DB	425	TTTCTACTGATACAGTTTCAAATTTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG	484
QY	506	GTAAAGAAAGAAAGGTAAAAACCGCTTAGAAGAACACGATAGAAATTTGAAGATATA	565
DB	485	CGAAAGAAAAGAGCTGAAGATTTTACTTTAAAGATACGATGATAAAGTACGCTGCATTCC	544
QY	566	AAAAAGA---AATACTATGGATTAATAAATCAAAGGTATTCCTGCAGTAGCTGCTAAT	622
DB	545	AAAAAGATCAAAGCAAGATATAAAGATGCAATGGCCATTGAAAGCTTCAGTTGTTAACT	604
QY	623	CAGGTTTGTCTCTCATCCAAACCAACTCTTATGTTGGTCAATTCCTTAAGTCAACTAGGTT	682
DB	605	TCCGT---GCTGATCATACAAAGATTTATGCTGGTGGATATGCTGTGAAATCTTAAATG	661
QY	683	TTAAAGAACATTAAGTATGATGTTTACTTAAAGGTTTAAAGTATCTTAAAGGACCTT	742
DB	662	ATTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGTTAAAGATATA	721
QY	743	ACTTACAAATGAACACTGAACTTTTATCTCAAGTGAATCCTGAGCGTATGTTTCATATGA	802
DB	722	TCCAACTTACATCTAAAGAAAGCAATTCCTAATTAATGAACGCTGATCATATTTTGTAGTAA	781
QY	803	CAAAACAAAGCAAGTTCTTAACGACCTTCACTA	834
DB	782	AATCAGATCCAAATGCGAAGATGCTGCATTA	813

US-08-781-986A-238
Sequence 238, Application US/08781986A
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP

US-10-329-624-238
; Sequence 238, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-10-329-624-238
Query Match 12.3%; Score 124; DB 18; Length 3775;
Best Local Similarity 50.7%; Pred. No. 1.5e-11;
Matches 412; Conservative 0; Mismatches 380; Indels 20; Gaps 4;
QY 27 AATTTTAGTGTAAATGCGCTTATTTGTTTAAATGCAACTGCGACATGCGAAATAA 86
DB 60 AATTAATAATGCTGTGTTTACGCTTCTCTCTTTTGTAGCAGATGTAGTGGAA 119
QY 87 TAGTTCAGTAAGTCAAGTAAAGAGTCAATCAAGATGGAGTTGAAATCAAGCAGCAAGA 146
DB 120 TTCAATAAACAATCATCTGATAACAAGATAAGGAACAACATTCATTAACATGCAT 179
QY 147 AGGTACTAGAAAGTACCTAAACACCCCTAAACGCTGTTGTTGTTTGTAGTATTCATTTGT 206
DB 180 GGGTACAACTGAAATTAAGGGAACCAAGCGCTGTTGTACGCTATATCAAGGTGCCAC 239
QY 207 TGATCGGTAGTGTGTTTGTAGTGTAAACCTGTTGGATAGCGGATGATAACAAAAAAA 266
DB 240 TGACGTCGCTGTATCTTTTAGGTGTTAAACCTGTTAGTGTGTTAGTATCATGGACAAAA 299
QY 267 TCGTATTATTAACCAATTAAGAGATAAAATGAAAAATACACTTCTGTAGGAACACGTAA 326
DB 300 ACCGAAATTCGATACATAAAAAATGATTTAAAAAGATCTAAGATTGTAGGTCAAGAAC 359

QY 327 GCAACCTAAGTACAGAAATCAGTAACCTAAACCGAGATTAAATTTTCTGCTGATATAA 386
DB 360 TGCACCTAAGTACAGAGAAATCTCTAAATTAACCGAGCTTAATTTGTCGGTCAAAAGT 419
QY 387 TAGACACAAAGGTATTTATTAAGAGCTTTAAATTAATAATGCTCTCCTACGATTGAACCTGAAAAAG 446
DB 420 TAGAAATGAAAAAGTTTACGATCAATTTATCTAAATTCGACC-----AACAG 466
QY 447 TTTCTGATGGAGATTATTAATGAAAAAT-ATTGATGCTTTTAAAAACAATTTCAAAAAGCTTTAG 505
DB 467 TTTCTCTGATACAGTTTCAAAATTCAAAGATACAACTAAGTTAATGGGAAAGCTTTAG 526
QY 506 GTAAAGAAAGAAAGGTAAAAAACGCTTAGAAGAACAGATGAAGAAAAATTTGAAGAATAA 565
DB 527 GGAAGAAAAAGAGCTGAAGATTTACTTTAAAAAGTAGATGAATAAAGTAGCTGCATTTCC 586
QY 566 AAAAAGA---AATAACTATGGATAAAAAATCAAAAGGTATTGCCCTGCAGTAGCTGCTAAAT 622
DB 587 AAAAAGATGCAAAAGCAAGATTAAGATGATGCGCATTTGAAAGCTTCAGTTGTTAACT 646
QY 623 CAGGTTTGTCTTCATCCAAAGCAACTCTTATGTTGTTCAATTCCTAAGTCAACTAGGTT 682
DB 647 TCCGT---GCTGATCATACAAAGATTTATGCTGCTGGATATGCTGGTGAATCTTAAATG 703
QY 683 TTAAGAAGCAATTAAGTATGATGTTACTAAGGTTTAAAGTAAAGTATCTTAAAGGACCTT 742
DB 704 ATTTAGGATTCAAAACGTAATAAAGACTTTACAAAACAAAGTTGTAATGGTAAAGATATA 763
QY 743 ACTTACAAATGAACACTGAAACTTTTATCTCAAGTGAATCCTGAGCGTATGTTTCAATAATGA 802
DB 764 TCCAACCTACATCTAAGAAAGCAATTCGATTAATGAACGCTGATCATATTTTGTAGTAA 823
QY 803 CAAACAAAGCAAGTTCTTAACGAACCTTCACTA 834
DB 824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855
RESULT 27
US-10-282-122A-10009
; Sequence 10009, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10009
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; US-10-282-122A-10009

Query Match      11.2%; Score 113.2; DB 17; Length 796;
Best Local Similarity 57.6%; Pred. No. 7e-10;
Matches 227; Conservative 0; Mismatches 158; Indels 9; Gaps 1

QY    191 TTGAGTATTCAATTTGGTGATCGGTGATTGCTTTTAGATGTTAAACCTGTTGGGATAGCGG 250
Db          |||||
QY    251 ATGNATACAAAAAAATCGTATTATTAAACCATTAAAGAGATAAAATTCGAAAATACACTT 310
Db          |||||
QY    62 ACATTAGAATATTATAATAAATGGGTAAATACAAAAACAACCCGAGTAAGATGTTGTAG 121
Db          |||||
QY    311 CTGTAGGAACAACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAACCTTAAACACAGATTTAA 370
Db          |||||
QY    122 ATGTGGGCACACGTCACACCAAACTTAGAAGAAATTAGCCGTTTAAACACGATTTAA 181
Db          |||||
QY    371 TTATTGCTGATAATAATAGACACAAAGGTATTATTAAAGACTTAAATAAAATGTCCTCTTA 430
Db          |||||
QY    182 TTATCACAGCTTCATTCCGTGGTAAAGCAATTTAAAAATGAAATTAGAACAAATTCACCAA 241
Db          |||||
QY    431 CGAT-----TCAACTGAAGAGTTTCGATGGAGATTATAATCGAAATATTGATGCTT 481
Db          |||||
QY    242 CAGTTATGTTTGATCCATCAACAAGCAATAACGATCATCTTTGCTGAAATGACAGAAACAT 301
Db          |||||
QY    482 TTAACCAATTTTCAAAAGCTTTTAGTAAAGAAGAAAGGTAAAAAAGCGTTTAGAAGAAC 541
Db          |||||
QY    302 TTAACCAATTTGCAAAAGCAGTTGGAAGAAGAGAGAGGTAAAAAGATTATTAGCTGATA 361
Db          |||||
QY    542 ACGATAAGAAAATTGAAGAATATAAAAAAGAAAT 575
Db          |||||
QY    362 TGGATAAGCAATTTCGCTGATGCAAAAGCAAAAT 395
Db          |||||

```

```

RESULT 28
US-10-282-122A-19870
; Sequence 19870, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```

	PRIOR APPLICATION NUMBER: 60/207,727	
	PRIOR FILING DATE: 2000-05-26	
	PRIOR APPLICATION NUMBER: 60/230,335	
	PRIOR FILING DATE: 2000-09-06	
	PRIOR APPLICATION NUMBER: 60/230,347	
	PRIOR FILING DATE: 2000-09-09	
	PRIOR APPLICATION NUMBER: 60/242,578	
	PRIOR FILING DATE: 2000-10-23	
	PRIOR APPLICATION NUMBER: 60/253,625	
	PRIOR FILING DATE: 2000-11-27	
	PRIOR APPLICATION NUMBER: 60/257,931	
	PRIOR FILING DATE: 2000-12-22	
	PRIOR APPLICATION NUMBER: 60/267,636	
	PRIOR FILING DATE: 2001-02-09	
	PRIOR APPLICATION NUMBER: 60/269,308	
	PRIOR FILING DATE: 2001-02-16	
	Remaining Prior Application data removed - See File Wrapper or PALM.	
	NUMBER OF SEQ ID NOS: 78614	
	SOFTWARE: PatentIn version 3.1	
	SEQ ID NO 19870	
	LENGTH: 900	
	TYPE: DNA	
	ORGANISM: Enterobacter cloacae	
	US-10-282-122A-19870	
	Query Match 10.2%; Score 103.2; DB 17; Length 900;	
	Best Local Similarity 52.9%; Pred.No.3.6e-08;	
	Matches 222; Conservative 0; Mismatches 198; Indels 0; Gaps 0	
Qy	137 AGCAGGAAGGTACTACGNAAGTACCCTAAACACCCCTAAACGGTGTTGTTGTCCTTGAGT	196
Dd	74 AGGACGAAACCGGCACGTTTACACTCGATAAACGCCACAACGGATTGGTGCTGGAAAC	133
Qy	197 ATTCATTTGTTGATCGCGTTAGTTCCTTTAGTAGTTTAAACCTGTTGGGATAGCGGATGATA	256
Dd	134 TCTCGTTCCGATGCGCTGSCCGCGTGGACGTCAGCCCAGTCGTTATGCGCACGATA	193
Qy	257 ACAAAAAAATCGTATTATTAACCNNTAAGAGATAAAATTTGGAATAATACACTTCTGTAG	316
Dd	194 ACGATGCAAAACGCGATCTCTGCCGAAGTGCGTGGCGCACTGAAAACCGTGGCAGTCCGTCG	253
Qy	317 GAACACGTAGCAACCTTAACCTTTAGAAGAAATCAGTAAACTTTAAACCGATTTAAATTAATG	376
Dd	254 GAACCGCGCGCAGCGAGCTGGAAGCCATTGTGCGCTGNAACGACCTGATCATTTG	313
Qy	377 CTGATAATAATAGACACAAAGGTAATTTATAAAGACTTTAAATAAANTGCTCTCAGATTG	436
Dd	314 CCGACAGCAGTCGCGCATCGCGGGATTACACTGCGCTTCGACGAAATCGCGCGGTACTGC	373
Qy	437 AACGTAAAGTTTTCGATCGGAGTATTAAATGAAATATTGATGCTTTTAAACCAATTTCAA	496
Dd	374 TGCTTAAAGTCCCGCAACGAAACCTACGCTGAAATTTGCCATCAGCGGCTATCATCGGGC	433
Qy	497 AAGCTTTAGGTAAAGAAGAAAGGCTAAAAACCGTTTAGAAGAAACAGATAGAATAATTG	556
b	434 AAGTGTAGGTATAAATACGACAGATGACGGCGCTCTCGAAACAAATAAAGAGAAGATGG	493

RESULT 29
US-09-815-242-6372
; Sequence '6372, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L. W.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 6372
LENGTH: 909
TYPE: DNA
ORGANISM: Escherichia coli
FEATURES:
NAME/KEY: CDS
LOCATION: (1)...(909)
US-09-815-342-6372

Query Match 9.4%; Score 95.2; DB 9; Length 909;
Best Local Similarity 51.7%; Pred. No. 8.4e-07;
Matches 217; Conservative 0; Mismatches 203; Indels 0; Gaps 0;
Qy 137 AGCAGAAAGAGGACTACGAAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGTAGT 196
Db 80 AGGACGAACACGGCAGCTTTACCTCGAAGAAACGCCACACGGATTGTTGCTGGAAAC 139
Qy 197 ATTCATTTTGTGTCGTTAGTTCGTTTGTAGATGTTAAACCTTGGGATAGCGGATGATA 256
Db 140 TCTCGTTCGCGATGCGCTGCGCGGCGGAGCTCATCCCGATCGGTATTCGCGACGATA 199
Qy 257 ACAAAAGAAATCGTATTATTAACCCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 316
Db 200 AGATGCAAAACGATCTCTGCCGAGTGGCGACCTGAAACCGTGGCAGTCCGTCG 259
Qy 317 GAACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAATCTTAACACGATTTAATTTG 376
Db 260 GAAACGCGCGCAGCGGAGCTGGAAGCCATTGCGGCTCTGAACACGACCTGATCATTTG 319
Qy 377 CTGATAATAATAGACACAAAGTATTTATAAGACTTAAATAAATTTGCTCTTACGATTG 436
Db 320 CCGACAGCAGTCGCCATCGCGGGGTTTACATCGCCTTGCAGCAATCGCGCGGTACTGC 379
Qy 437 AACTGAAAGTTTCGATGAGATTATAATGAAATATTGATGCTTTTAAACAAATTTCAA 496
Db 380 TGCTTAAGTCCGCAACGAACTACGCTGAAATTTGCAATCTCGCGCTATCATCGCGG 439
Qy 497 AAGCTTTAGGTAAGAAGAGAGGTAAGAAACCGTTAGAGAAACACGATTAAGAAATTTG 556
Db 440 AAATGGTGGGTAAAGAGCAGAGATGCGAGCAGCTCTGGAACAAACATTAAGAGAGGATGG 499

RESULT 30

US-10-282-122A-20621
Sequence 20621, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel

APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20621
LENGTH: 909
TYPE: DNA
ORGANISM: Escherichia coli
US-10-282-122A-20621

Query Match 9.4%; Score 95.2; DB 17; Length 909;
Best Local Similarity 51.7%; Pred. No. 8.4e-07;
Matches 217; Conservative 0; Mismatches 203; Indels 0; Gaps 0;
Qy 137 AGCAGAAAGAGGACTACGAAAGTACCTAAACACCCCTAAACGTTGTTGTTCTTGTAGT 196
Db 80 AGGACGAACACGGCAGCTTTACCTCGAAGAAACGCCACACGGATTGTTGCTGGAAAC 139
Qy 197 ATTCATTTTGTGTCGTTAGTTCGTTTGTAGATGTTAAACCTTGGGATAGCGGATGATA 256
Db 140 TCTCGTTCGCGATGCGCTGCGCGGCGGAGCTCATCCCGATCGGTATTCGCGACGATA 199
Qy 257 ACAAAAGAAATCGTATTATTAACCCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 316
Db 200 AGATGCAAAACGATCTCTGCCGAGTGGCGACCTGAAACCGTGGCAGTCCGTCG 259
Qy 317 GAACACGTAAGCAACCTTAACCTTAGAAGAAATCAGTAAATCTTAACACGATTTAATTTG 376
Db 260 GAAACGCGCGCAGCGGAGCTGGAAGCCATTGCGGCTCTGAACACGACCTGATCATTTG 319
Qy 377 CTGATAATAATAGACACAAAGTATTTATAAGACTTAAATAAATTTGCTCTTACGATTG 436
Db 320 CCGACAGCAGTCGCCATCGCGGGGTTTACATCGCCTTGCAGCAATCGCGCGGTACTGC 379
Qy 437 AACTGAAAGTTTCGATGAGATTATAATGAAATATTGATGCTTTTAAACAAATTTCAA 496
Db 380 TGCTTAAGTCCGCAACGAACTACGCTGAAATTTGCAATCTCGCGCTATCATCGCGG 439
Qy 497 AAGCTTTAGGTAAGAAGAGAGGTAAGAAACCGTTAGAGAAACACGATTAAGAAATTTG 556
Db 440 AAATGGTGGGTAAAGAGCAGAGATGCGAGCAGCTCTGGAACAAACATTAAGAGAGGATGG 499

RESULT 31

US-10-893-671-14/c

```

; Sequence 14, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al.
; TITLE OF INVENTION: NIMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 14
; LENGTH: 10244
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-10-893-671-14

Query Match          9.4%; Score 95.2; DB 21; Length 10244;
Best Local Similarity 51.7%; Pred. No. 1.6e-06;
Matches 217; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

Qy      137  AGCACGAAGAAGGTACTACGAAGTACTTAAACACCCCTAAACGTTGTGTGTTCTTGAGT 196
Db      7731  AGGACGAACACGGCACCTTTACACTCGAAAAAACGCCACACGGATTGTGTGCTCGAAC 7672

Qy      197  ATTCATTGTTGTATCGCTTGTGCTTTAGATGTTAAACCTGTTGGGATAGCGATGATA 256
Db      7671  TCTCGTTGCGCGATGCGCTGCGCCCGGTGGACGTCATCCCGATCGGTATTGCGCAGCATA 7612

Qy      257  ACAAAAAAATTCGPTATTATTAAACCATTAAGAGATAAAATTTGAAAAATACACTTCTGTAG 316
Db      7611  ACGATGCAAAACGGATCCTGCCCGAAGTGCGTGGCGACCTGAAACCGTGGCAGTCGCTCG 7552

Qy      317  GAACACGTGAGCAACCTTAACCTTTAGAGAAATCAGTAAACCTTAAACACAGATTTAATTATG 376
Db      7551  GAACGCGCGCAGCCGAGCTGGAAGCCATTGCGCGCTCTGAAACACACAGACTGATCATTTG 7492

Qy      377  CTGATAATATAGACACAAAGGTATTATTAAAGACTTTTAAATAAAATTCCTCTACGATTG 436
Db      7491  CCGACAGCAGTCGCGCATTCGCGGGGTTTACATCGCCTTTCGAGCAATTCGCGCGGTACTGC 7432

Qy      437  AACTGAAAAAGTTTCGATGGAGATTATATGAAAAATATTGATGCTTTTAAAAACCAATTTCAA 496
Db      7431  TGCCTTAAGTCCCGCAACGNAACCTACGCTGAAAAATTTGCAATCTCGCGCTATCATCGCG 7372

Qy      497  AAGCTTTTAGTAAAGAAGAAGAGGTAAAAAACGCTTAGAAGAACACAGATAAGAAATTTG 556
Db      7371  AAATGCTGTGGTTAAAGACGAGAGATGCAAGGACCGTCTGGAAACAACAATAAGAGAGATGG 7312

```

RESULT 32
US-10-282-122A-17083
; Sequence 17083, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohleen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A

```

/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 17083
/ LENGTH: 963
/ TYPE: DNA
/ ORGANISM: Clostridium difficile
US-10-282-122A-17083

Query Match          7.8%; Score 79; DB 17; Length 963;
Best Local Similarity 47.1%; Pred. No. 0.00049;
Matches 312; Conservative 0; Mismatches 345; Indels 6; Gaps 2;

Qy 27 AATTTTAAGTGTAATTGGCTTATTGTTGTTTAAATTGCAACTGCGACATGTGGAATAAA 86
Db 18 AATAGTAGCAGCAATAGCTATTATAGGATTAGTAGCAGAGTGTTCATTAGTGGAGTAA 77

Qy 87 TAGTTCAAAGTAACCTCAAGTAAGAGAGTCATCAAAGATGGAGTTGAAATCAAGCAAGA 146
Db 78 AAAGAATGAATCAAAACACATCAGAAAATTCAAATATAACCAATAAGATAACTCATAACTT 137

Qy 147 AGGTACTACGAAAGTACCTAAACACCCCTAAACCGTGTGTTGTTCTTGAGTATTCATTGT 206
Db 138 GCGAGAAACAGATGTAAATTTAAATCCTAAAAAAGTAGTAGTATTTTGATTATTCGGCTTT 197

Qy 207 TGATGGCTTAGTTCCTTTAGACTGTAAACCTGTTGGGATAGCGGATGATAACAAAAAAA 266
Db 198 AGATACAATGGATGCATTAGGTGT---AGCTAGAAATTTAGTAGGACTTCCAAAAGCGAG 254

Qy 267 TCGTATTATTAAACCATTAAGAGATAAAATTTGAAAAATACACTTCTGTAGSAAACAGTAA 326
Db 255 TTTACCAGCATCTTTAGAAAAATATAAGGATAGAAGTATACAGATTTAGGAGGCTTAAA 314

Qy 327 GCAACCTPACTTAGAAGAAATCAGTAAACCTTAAACCGAGATTAAATTATTCGTGATAATA 386
Db 315 AGAACCAAGATTTAGAAGGTATAAGTCGCGAAATCCAGATTTTAATTATAATAATGGAAG 374

Qy 387 TAGACACAAAGGTATTTATAAAGACTTTAAATAAAATTTGCTCCCTACGATTGCACTGAAAAG 446
Db 375 ACAAGAAGATTTTTATGAGCAATTTGTCAAAAATAGCACCACAAATAAGTACAGTAAAGA 434

Qy 447 TTTTCGATGGAGATTATATGAAAAATATTGATGCTTTTAAAAACAATTTCAAAGCTTTAGG 506
Db 435 TGATAGAAGTACTTAGAATCTGTTAAAAATATATAGACAAATAGCAAAATATTATTTGG 494

Qy 507 TAAAGAAGAAGAGGTAAAAACCGCTTAGAAGAAACAGATAGAAGAAATTTGAAGAAATATA 566
Db 495 AGTAGAAGAAAAAGCAAAATCAAGAATTTTAGTAAAAATTTGAAAAGAAAAATAGAAACCTTTGAA 554

Qy 567 AAAAGAAATTAACATAGTAAAAATCAAAAGGTATTCCTGCAGTGTCTTAATCAGG 626
Db 555 TAAAAAAGTAGTACA---GATAAAAATTTAAATGCTTTTAACATATATGTTAAATGAAGGTAA 611

```



```
QY 494 CAARAGCTTTAGGTAAGAGAGAGAGCTTAAARACGCTTAGAAGAACACGATAGAAA 553
Db 184 TAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 125
QY 554 TTGAAGAAATATAAAAAAGAAATAAATACTATGGATAAAAAATCAAAAGGTATTG 603
Db 124 ATAAAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATTATAG 75

RESULT 35
US-10-021-323-6774
; Sequence 6774, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 6774
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(478)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3828-010-Q1-N6-E1
US-10-021-323-6774

Query Match 7.7%; Score 77.8; DB 19; Length 478;
Best Local Similarity 49.3%; Pred. No. 0.00065;
Matches 202; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

QY 180 TCTGTGTTCTTGTAGTATTCATTTGTTGATGCTGTTAGTTGCTTTAGATGTTAAACCTGT 239
Db 6 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 65

QY 240 TGGGATAGCGGATGATACAAACAAAAATCGTATTTATTAACCATTAAGAGATAAAATTGG 299
Db 66 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 125

QY 300 AAAATACACTTCTGTAGGAACAGCTAAGCAACCTAATCTTAGAAGAAATCAGTAAACTTAA 359
Db 126 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 185

QY 360 ACCAGATTTTATTTGCTGATATAATAGACACAAAGGTATTTATTAAGAGCTTAAATAA 419
Db 186 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 245

QY 420 AATTTGCTCTACGATTTGAACCTGAAAGTTTTCGATGGAGATTTAATAATGAAATATTGATGC 479
Db 246 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 305

QY 480 TTTTAAACAAATTTCAAAAGCTTTAGGTAAAGAAAGAGGTAAAAAACGCTTAGAAGA 539
Db 306 GGAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAAGA 365

QY 540 ACAGATACAAATTTGAGCAATATAAAGAAATTAACCTATGATATAAA 569
Db 366 AAAGGAGGGGGAATAATGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 415

RESULT 36
US-10-282-122A-33236
; Sequence 33236, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33236
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
US-10-282-122A-33236

Query Match 7.5%; Score 75.8; DB 17; Length 891;
Best Local Similarity 47.7%; Pred. No. 0.0017;
Matches 221; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY 128 TTGAATCAAGCACGACGAAGGTACTACGAAGTACCTTAACACACCTTAACCGTGTGTG 187
Db 53 TGCATCATGATGACGCCGACCAAGGTGTCATCTGCCAGACACCCCAAGCGCGTGTGG 112

QY 188 TTCTTTAGTATTTCATTTGTTGATGCTTAGTTGCTTTAGATGTTAAACCTGTGTGGATAG 247
Db 113 TACTGGAATTTTCGTTTCTCGACGGGCTTCTCGTGGGTGACACGGTTCGGGCGG 172

QY 248 CGGATGATAACAAAAAATAATCGTATTTATAACCAATTAAGAGATAAAATTTGAAATAACA 307
Db 173 CAGATGATGGCGACGCCAGTCGTGCTAGTCTGCCCAAGGTGCGCAAGCGCGTGAATGC 232

QY 308 CTTCTGTAGGAACACGTAAGCAACCTTAAGAGAAATCAGTAAACTTAAACCAAGATT 367
Db 233 AGTCGTGGGGCTGCGCTCGCAACCAATATCGAAGTATCGCACGGCTCAAGCGCGGACC 292

QY 368 TAATTTATTCGTCTGATTAATAGACACAAAGGTATTTATAAAGACTTTAAATAAAATTTGCTC 427
Db 293 TGAATCATGCTGACCTCGGTCTGTCATCAGGCTCTTTATTAACGACCTGGCCAGCTGCTC 352

QY 428 CTACGATTGAACCTGAAAAAGTTTCGATGGAGATTATAATGAAAAATATTGATGCTTTTAAAA 487
```



```

353 CGACGCGTGATGTTGCCGTCA CGCGCGCGAGGACTATCAGGCGCAGCCTGAAATCTGCCGGGC 412
488 CAAATTTCAAAGAGCTTTAGGTAAAGAAAGAAAGGTAAGAAAAACGCTTAGAAGAAACACGATA 547
413 TGATCGCGATCGCGCTGGGCAAGAGCCCGGAGATCGACGGCGCGAATCGCAGAAAACCGTC 472
548 AGAAAAATTGAAGATATATAAAAAAGAAATAAATCTATGGATAAAAA 590
473 AGCACCTGAAAAACCGTTGCCGAGCAGATTCTCTGCCGACAGCAA 515

RESULT 37
US-10-425-115-115706/c
; Sequence 115706, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 115706
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1121)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_37006C.1
US-10-425-115-115706

```

Query Match	7.5%	Score	75.6;	DB	20;	Length	1121;
Best Local Similarity	44.7%;	Prod.	No.	0.0019;			
Matches	288;	Conservative	0;	Mismatches	356;	Indels	0;
Gaps	0;						
Qy	255	TAA	CAAAAAA	AAATCGTATTTATTA	AAACCAATTAAGAGATAAAAT	TGAAATAACACTTCGT	314
Db	1089	TAAAAAT	AAAAAGAAAA	CAAAAAA	CAAAAAAATTA	AAAAA	1030
Qy	315	AGGAACACGT	AGCAACTTA	CTTAGAGAAATCAGTAA	AACTTAAAC	CAGATTTTAAT	374
Db	1029	AAATG	AAAAA	AAAGTAAAAACG	AAAAA	AAAAA	970
Qy	375	TGCTGAT	TAATATAGACACA	CAAAAGGTATTTATTA	AGACCTTAATAATAA	TGCTCTACGAT	434
Db	969	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	910
Qy	435	TGAACTG	AAAAAGTTTCGAT	TGGAGATTATTAAT	CGAAATATTGAT	TGCTTTTAAACAACTTC	494
Db	909	AAAA	CAAAAAA	AAATATGA	AAAAA	AAAAA	850
Qy	495	AAAAAGCTTTAGGT	TAAAGAAAGAGGTAAGG	TAAAAA	ACGCTTAGAAGAAC	ACGATAAGAAAAAT	554
Db	849	AAAAA	ATGTAAAAA	AAAAA	AAAAA	AAAAA	790
Qy	555	TGAAGAATAT	AAAAA	GAATATACTAT	CGATATAA	AAATCAAAAGGTATTGCTCGCATGAC	614
Db	789	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	730
Qy	615	TGCTAAATCAGGTTTGC	TGCTTGC	TCTCAACGAAC	TCTTTAT	TGTTGTCATTCCTCAAGTCA	674
Db	729	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	670
Qy	675	ACTAGGTTTTTAA	GAAGCATTTAAGT	GTGATGTTACT	TAAAGGGTTTTAAGT	TAAGTATCTTAA	734

669 AAAAAAAAAAAAAAAAAAATAAAGTAATAAAAAAAAAACAAAAAATAATAATAA 610

735 AGGACCTTACTTACAAATGAACACTGAAACCTTTATCTCAAGTGAATCCTGAGCGTATGTT 794

609 AAACAATATTATTAAAAAAAAAAAAAAAAAATAACAATAAAAAAAAAACAAAAACAATAAATAATA 550

795 CATTAATGACAAACAAAGCAAGTTCTTAAGGAACCTTCTAATAAGAACTAGAAAAAGATCC 854

549 AAAAACCATTAATAAAAAAAAAAAAAAAAAATAATAACAAAAACACAAAAAATAAATAAATAATNC 490

855 TGTATGGAAGAAATTAACACGCTGTGAAAAAATCAACGTGTTGATA 898

489 ACTACATAATAGATCAAAAAAAAAACCAAAACATACAAAACTGATA 446

RESULT 38

US-10-425-115-120013/c

; Sequence 120013, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 120013

; LENGTH: 1062

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(1062)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MFT4577_40935C.1

US-10-425-115-120013

	Query Match	7.4%	Score 74.2;	DB 20;	Length 1062;
	Best Local Similarity	51.0%;	Pred. No. 0.0033;		
	Matches 175;	Conservative 0;	Mismatches 168;	Indels 0;	Gaps 0;
QY	254	ATAACAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTTGGAAAAACACATTCG	313		
DB	919	AAGA	860		
QY	314	TAGGAACACGTAAAGCAACTTAACCTTAGAGAAATTCAGTAAACTTTAAACCAAGATTTAATTA	373		
DB	859	AAAAAAAACAACAAATTA	800		
QY	374	TTGCTGATAATAATAGACACAAAGGTATTATTAAAGACTTTAAATAAATTTGCTCCTACGA	433		
DB	799	AAATAA	740		
QY	434	TTGAACCTGAAGTTTCGATCGAGATTATTAATGAAATATTGATGCTTTTAAACCAATTT	493		
DB	739	AAATAA	680		
QY	494	CAAAAGCTTTAGGTAAAGAAAGAGAGGTAAAAAACGCTTTAGAGAACAACGATAAGAAAA	553		
DB	679	AAAAAAAAAAGAAAAAAAAAAAAAAAAAATAAATAAAAAAAAAAAAAAAAAAAAAA	620		
QY	554	TTGAAGAAATATAAAAAAGAAATACACTATGATTAATAAATCAAAA	596		
DB	619	AAAAATAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAA	577		

RESULT 39
US-10-424-599-102083

```

; Sequence 102083, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102083
; LENGTH: 1214
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1214)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63198C.1
; US-10-424-599-102083

```

	Query Match	7.3%;	Score 74;	DB 18;	Length 1214;
	Best Local Similarity	48.7%;	Pred. No. 0.0037;		
	Matches 167;	Conservative 0;	Mismatches 176;	Indels 0;	Gaps 0;
Qy	254	ATACAAAAAAATCGTATTATTATTAACCACTTTAGAGATAAAATTTGGAAAAATACACTTCTG	313		
Db	860	ATAAAAAAATAAAAAAATAAAAAAACAATATAATAATAAAAAAATAAAAAATAAA	919		
Qy	314	TAGGAACACGTAAGCAACCTTAACCTTTAGAGAAAAATCAGTAAACCTTAAACCGAGTTTAAATTA	373		
Db	920	AAAAAAGGAATACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	979		
Qy	374	TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTTGCTCTTACGA	433		
Db	980	AANAATAAAAAAATAAAAAAANAANAANAANAANAANAANAANAANAANAANAANA	1039		
Qy	434	TTGAACTGAAAAGTTTCGATGGAGATTATATGAAATATTTGATGCTTTTAAACCAATTT	493		
Db	1040	AAAAAATAAAAAAATAAAAAAANAANAANAANAANAANAANAANAANAANAANAANA	1099		
Qy	494	CAAAAGCTTTTAGGTAGGAAGCAAGGTAAAAAACCGCTTAGAAGAACACGATAAGAAAA	553		
Db	1100	AAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA	1159		
Qy	554	TTGAAGATATATAAAAAAATAAATACTATGGATAAAAAATCAAAA	596		
Db	1160	NNNAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	1202		

```

RESULT 40
US-10-425-115-172717/c
; Sequence 172717, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 172717
; LENGTH: 1243
; TYPE: DNA
; ORGANISM: Zea mays

```

```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1243)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_89100C.1
US-10-425-115-172717

Query Match      7.3%; Score 74; DB 20; Length 1243;
Best Local Similarity 43.7%; Pred. No. 0.0037;
Matches 330; Conservative 0; Mismatches 422; Indels 3; Gaps 1;

Qy 251 ATGATACAAAAAATCGTATTATTAAACCATTAAAGAGATAAAATTTGGAAAAATACACTT 310
Db 1230 AANAAAAAANAAAAANAAAAAACAACAAAAANAAAAANAAAAANAAAAANAAAAA 1171

Qy 311 CTGTAGGAACACGTAAGCAACCTAACTTAGAGAAATCAGTAAACCTTTAAACCCAGATTAA 370
Db 1170 AAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1111

Qy 371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAGACTTATAATAAATTTGCTCTTA 430
Db 1110 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1051

Qy 431 CGATTGAACCTGAAAAGTTTCGATGGAGATTTATATGAAATATTTGATGCTTTTAAACAA 490
Db 1050 ATAAAAAATTAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 991

Qy 491 TTTCAAAGCTTTTAGGTAAAGAGAAGAGTAAAAAAACGCTTAGAAGAACACCGATAAGA 550
Db 990 ATAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 931

Qy 551 AAATTGAAGAATATAAAAAAGAAATACTATGGAATAAAAAATCAAAAGGTATTTGCTCGAG 610
Db 930 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 871

Qy 611 TAGCTGTAATCAGGTTGCTCTCATCCAAGCACTCTATGTTGGTCAATTTCTTAA 670
Db 870 TAAAAAANAANAANAATTAANAATAAAAAATAAAAAANAANAANAANAANAANAATATAGNA 811

Qy 671 GTCACCTAGGTTTTAAAGAGACATTAAGTGATGATGTTACTTAAAGGTTTAAGTAAGTATC 730
Db 810 AAAAAATAAATATAAATAATAAATAAANAANAANAANAANAANAANAANAANAANAANA 751

Qy 731 TTAAGGACCTTACTTACAAATGAACACTGAAACCTTTTATCTCAAGTGAATCCTGAGCGTA 790
Db 750 AATTAAAAAATAAGAAACAAGAAANAANAATAAAAAAATTAANAANAANAANAANAANA 691

Qy 791 TGTTCATAATGACAAACAAGCAAGTCTTAACGAACCTTCTCTAAAAGACTAGAAAAAG 850
Db 690 AAT--NAAATTAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 634

Qy 851 ATCTGTATGGAAGAAATTAACCGCTGTGAANAATCAACGTTGTTGATATTTTAGACCGTG 910
Db 633 ATCTATATAAATAAANAANAAGGAANAATAAANAANAANAAGGAATTTAAAAANAANAAGNA 574

Qy 911 ACTTATGGCAGAGTACGCTGGTTTAAATTTCTTCAGAGAAATGCGCAAGAACTTGTTG 970
Db 573 ACAAGAAAAAATTAANAANAANAAGNAATACACTTATAAATAAANAANAANAATATTANAATTA 514

Qy 971 AATTATCTTACAGAAAGATAGTAAAAAAGATAATAAG 1005
Db 513 AATATAAATAATANATAACCAAAACAAAAATAAG 479

```

Search completed: August 26, 2005, 12:35:16
Job time : 2126 secs

This Page Blank (unspliced)